

# nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

## DOUBLE PROTECTION

*Dissecting the cell-senescence  
and crisis barriers that block  
the road to cancer* **PAGE 492**

### BIOMEDICAL RESEARCH

#### LONDON CALLING

*Behind the scenes at the  
Francis Crick Institute*

**PAGE 406**

### GENE EDITING

#### THE CRISPR DEBATE

*Why democracy and  
pluralism must prevail*

**PAGES 413 & 415**

### ECOLOGY

#### CONSERVATION ON ICE

*Antarctic biodiversity highly  
vulnerable to human activity*

**PAGE 431**

**NATURE.COM/NATURE**

25 June 2015 £10

Vol. 522, No. 7557

# THIS WEEK

## EDITORIALS

**FROZEN** Antarctica is a distant but not a remote place **p.392**

**WORLD VIEW** My former boss Tim Hunt is no sexist **p.393**



**SOUTHPAW** Bipedal kangaroos like to lead with their left **p.395**

## Hope from the Pope

*The Vatican has produced a timely and valuable warning on the threat of climate change that will reach a wide audience.*

Relations between the Catholic Church and science have long been ambivalent. Famously, it took the Vatican 359 years to formally concede that Earth moves around the Sun: only in 1992 was Galileo Galilei, who proclaimed that fact in 1633, officially ‘rehabilitated’ by Pope John Paul II.

It is gratifying that Catholic officials have moved on. Pope Francis’s much-anticipated encyclical letter on climate and the environment, released last week with great fanfare, is of course pastoral in tone and intent. And yet, the papal letter testifies to the respect and honest understanding that the Argentinian Pope and his cardinals have for science.

“Science is the best tool by which we can listen to the cry of the Earth,” Cardinal Peter Turkson, president of the Pontifical Council for Justice and Peace, said on releasing the 192-page letter, drafted in consultation with leading climate experts. While calling on leaders to step up efforts to reduce global greenhouse emissions (which the encyclical acknowledges to be the main cause of warming), the Pope warns in strong language that unabated climate change will cause severe disruptions to ecosystems and human societies. His letter to the world’s more than one billion Catholics repeatedly stresses the moral imperative for action on global warming, poverty and overexploitation of natural resources.

With international climate negotiations entering a decisive phase, the Vatican has chosen the right time to weigh in. Combined with a pledge earlier this month from the G7 leading industrialized nations to phase out fossil fuels, hopes are increasing that the next round of global climate negotiations, in Paris in December, might reach a meaningful deal.

Scientists and political leaders in favour of climate action have rightly expressed gratitude and admiration for Pope Francis’s brave move. In doing so, President Barack Obama said that the United States must be a leader in efforts to cut carbon pollution and protect the environment. Other nations must follow suit. It will take more than spiritual aid to set the course for a clean-energy future, but the political weight of the Pope’s appeal to the moral conscience and reason of millions of people must not be underestimated.

The fact that the Vatican has sided with those who criticize the excesses of capitalism and consumerism — systems that the Pope labels in the encyclical as culprits in the world’s ecological crisis — could fuel suspicion in some quarters. But his plea for moderation and sustainability should not be dismissed as naive or utopian. The long-held view that continuous economic growth produces perpetual happiness and well-being is increasingly being questioned, including by forward-looking economists and social psychologists. And the financial crisis of recent years has produced ample evidence of the risks of runaway capitalism.

Wisely, Pope Francis has framed the issue of global change along the lines of hope and opportunity, rather than of doom and gloom. Why, he asks, should humankind take the risk of dangerously altering the planet’s climate and biosphere when there are viable alternatives?

The papal calls to end poverty and share the world’s ecological space in a fair way are objectives that mirror the United Nations’ Sustainable

Development Goals, to be released in September. The Pope’s letter adds an important facet to the discussion: it is not merely conceivable to secure a sound future for human civilization without relying on coal, oil and gas — it is a prerequisite.

Given the Pope’s moral authority and sky-rocketing popularity — not just among Catholics — his words might travel farther than sober scientific reports by bodies such as the Intergovernmental Panel on Climate Change (IPCC). Those tomes tend to brim with jargon and are read by relatively few people. And unlike the IPCC, which sometimes struggles with public outreach, the Catholic Church has a superb means of spreading the word every Sunday. A global campaign that the church plans to run in coming months could step up the urgency for the December climate talks. To build useful momentum, climate policies must be top of the agenda when Pope Francis visits the White House and addresses the UN General Assembly in September.

Meanwhile, within the encyclical are other issues that are important for science. Remarkably, the Pope concedes that intervening with plants and animals is “permissible when it pertains to the necessities of human life”. The clause signals a rethink of the Catholic Church’s attitude to genetic engineering and animal research. Alas, he remained silent on issues of contraception. With a world population heading towards a possible 10 billion, the importance of family planning is clear. The Vatican has been brave on climate change. If it is serious about the fate of the planet and the welfare of its inhabitants, then it must be braver still on the issue of contraception. ■

**“It will take more than spiritual aid to set the course for a clean-energy future.”**

## Data overprotection

*Draft European rules governing privacy threaten to hamper medical research.*

When officials at the European Commission proposed new data-protection rules in 2012, the prospects for science looked good. Three years on, that optimism has been replaced by concern. The best that researchers can hope for from the rules now, it seems, is that they do not make things worse. For, as politicians continue to try to protect the individual in a digital world, they risk inflicting major long-term damage on the research environment.

The rules aim to update the 1995 data-protection regulations to reflect the reality of the digital age, in which information about individuals is increasingly being used as a commodity.



A pan-European law on how data could be sourced, stored and used would, in theory, be good for research. Greater harmonization could smooth the difficulties that scientists face when they try to pool analysis of genomic data and tissue samples across national borders. Such sharing could help to organize powerful trials with large numbers of participants. But it is held back at present because different European countries have their own rules on issues such as informed consent, or on how to anonymize or pseudonymize data.

Unfortunately, as a result of their reworking, the draft rules threaten to place further restrictions on such studies, by tightening the way that samples and data can be used. The situation is so perilous that researchers met in Brussels last week for a 'Day of Action on Data for Health and Science', to explain to policy-makers "the reasons why personal data is necessary to scientific research, including medical research".

The reasons might sound obvious to scientists, and the need to explain them to politicians may seem distressing. But this is no anti-science agenda at play here. No one had intended to make medical research more difficult. The trouble is that such research is only a small part of the legislative issue. And discussions all take place in ministries of justice and other places where grasp of the mysterious world of science is relatively weak.

There is hope, in other words, that the feared damage can be avoided when the draft rules are finalized, possibly by the end of this year, as long as enough people make enough fuss to attract enough attention. To their credit, science funders and lobby groups across Europe have been doing just that.

The regulations proposed in 2012 by the commission allow the use of personal data only after an individual has given "specific, informed and explicit" consent. It is, after all, directed primarily at potential misuse of data gathered at websites or apps such as Facebook. But it made exceptions for health research because it recognized that such

work depends on broad models of consent in which participants can agree for their data to be used in a variety of research studies, and that — subject to suitable safeguards — the data can be stored for future use.

Parliamentary amendments agreed last year changed the emphasis, and in doing so removed the exemption for research. If accepted, the amendments would require consent from participants for each new study. That would impede research. It is something that patient groups say they do not want if they donate data or tissue to a trusted party, such as established biobanks. And the evidence suggests that participants are less likely to give re-consent the more frequently they are asked for it.

Talks on the draft rules were due to restart this week, with triologue negotiations scheduled between the European Commission, Parliament and Council. The council, at least, seems to be aware of the unwitting threat to research and wants to do something about it. Its statement on the issue, published last week, indicates that it will argue for the research exemption to be maintained.

Why should science be seen as a special case? A rigorous governance and regulatory framework for biomedical research already exists and operates under national and international laws. As this journal has chronicled, the 'big data' era poses new challenges, to privacy for example. But from a scientific standpoint, it makes more sense to update these existing rules as required than for valuable research to be threatened by overarching laws aimed at a different problem.

The European data-protection drive is built on the principle of "one continent, one law, one regulation". But even the most committed believers may have to accept that biomedical research must remain an exception — and be granted an exemption — if it is to continue to thrive. ■

# Life under the ice

*Antarctica's apparent barrenness hides an abundance of living organisms.*

Many of our views of the continent of Antarctica are influenced by the words of the great polar explorers. Robert Falcon Scott, a man ultimately drawn to his death there, famously wrote of the Antarctic, "Great God, this is an awful place". Present-day adventurers, from those who follow in the footsteps of Scott to the interior to tourists skirting around the edge, also trade on the sheer other-worldliness of Earth's southern cap. The more remote Antarctica is, the more their sense of achievement. Head there by ship and — if you can stand the swell — you can peer over the side to watch the seawater change colour at the convergence point where cold currents circling the continent push up against the warmer waters of the southern Pacific Ocean.

To stand on Antarctica is truly like nothing else on Earth. On a still day, it is as if time itself has frozen. Senses, the essential guides to life on our planet, are almost redundant. There is no movement, no sound and no smell. Scientists who spend time there have been known to take curry powder on their field expeditions, to prepare spicy food to compensate for the lack of other stimulation. With centuries-old snow beneath your feet and only the occasional rumble of distant ice cliffs collapsing into the ocean, it is easy to imagine Antarctica — and those who spend time there — as isolated from the rest of the world and its life.

Look in a different direction, however, and the illusion fades. Scott knew this. "As one looks across the barren stretches of the pack, it is

sometimes difficult to realise what teeming life exists immediately beneath its surface," he wrote. "Beneath the placid ice floes and under the calm water pools the old universal warfare is raging incessantly in the struggle for existence." There is life in Antarctica; you just need to know where to look.

For a sense of the teeming life that exists in the white continent, take a look at the Review article that we publish this week on page 431. Yes, there are only two species of flowering plants in Antarctica, but nobody said that the place had to look pretty. Think small. More than 200 species of lichen and 100 species of moss cling to the coastline and to patches of exposed rock. The packed snow that rests on the Antarctic bedrock is riddled with water-filled holes, and these subglacial lakes support the highest diversity of viruses of any known aquatic system.

Still not impressed? Head offshore — not too far — and the Antarctic waters thrive with so much life that they are considered to harbour more biodiversity than the celebrated Galapagos Islands. Throughout the Southern Ocean, researchers have discovered more than 8,000 marine species, most of which scrape a living from the seabed. From such a rich base, a network of life spreads upwards and outwards until it culminates in the region's iconic seals, penguins, whales and albatrosses.

The point of this continental audit is threefold. Antarctica is alive — and therefore dynamic. Some change is bad and some good, but most points in the same direction: that Antarctica's isolation from the rest of the world — biological, physical and cultural — is weakening. And in a world that increasingly looks to value from the natural environment, the Review article sums up the place nicely. "Antarctic ecosystems provide several services essential to planetary stability." Awful or not, Antarctica is closer to home than many realize. ■

➔ **NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunqv](http://go.nature.com/xhunqv)



## Judge by actions, not words

*Sexist comments made by my former boss Tim Hunt are not an indication that he is biased against women, argues Alessia Errico.*

From 2004 until December 2010, I worked as a scientist in the London laboratory of Tim Hunt. Given the media frenzy over his comments this month about women scientists, and the way he has been portrayed as sexist, you might think that I had a miserable time. Not at all; it was a great experience. He was an inspiring and supportive mentor. Not once did I feel that he treated me differently because I was a woman.

The anger that many people in research feel at Tim's inappropriate and indefensible comments — that women and men in the lab fall in love and that women cry when criticized — is understandable. But the reaction to them, and their impact on his career, has been disproportionate. Tim has apologized for his remarks, and I, for one, accept that apology. Others should do the same, and he should be allowed to continue his work.

Tim says that his comments were intended as a joke. I believe him, because I was once on the receiving end of one of his public jokes — and his attempt at humour then was just as bad.

My early research was funded by a European consortium. Each year, Tim and I had to attend the consortium conference, where the group leaders would discuss the work. As we travelled to the first of these conferences in April 2005, Tim realized that he did not have a photo of me to include in his talk, which was that afternoon. He always travelled with a camera in his shirt pocket, so he took a photo then. During the talk, he introduced the picture to the audience with the line: "There is Alessia, wearing exactly the same jumper." Of course, everyone in the audience searched for me among themselves and, having verified that I was indeed wearing that jumper, had a good laugh.

I was so embarrassed. There I was, at the first meeting to discuss my research, and I had to start by explaining to new colleagues that, yes, I did own more than one set of clothes. Still, I was not offended. I knew he was joking, and it did provide a way of breaking the ice when I met people later.

The fear of many scientists, and of those organizations that have severed their ties with Tim, is that his judgement might be biased. I know him well and can say that he is not sexist. I once attended a 'Women in Science' event where his wife Mary was one of the speakers and where many of the issues that we have read about in the media were discussed. Tim was also in the audience, and I know that he engaged with the problem of discrimination against women scientists and supports efforts to tackle it.

I have seen discrimination and sexism in

science and in wider society. I have seen female colleagues talked about in negative ways when they left the lab to have children. The issue is a genuine one that demands urgent attention. But it is grossly unfair that Tim should be considered, and treated, as an emblem of this sexism or gender discrimination.

I treasure all scientific discussions with Tim; he always had an interesting point of view and would constructively challenge us all. When I delivered my first seminar at our institute, Tim discussed the slides with me several times to ensure that the logic behind the project and each experiment was clear. I gave several talks while I was in the lab and his contribution to my preparation was always useful and positive. Although very busy, whenever he could he visited the lab, and would pop by to say hi and ask how things were going.

He always supported my work, even when my research project took a different direction from the lab's main expertise. For example, he arranged for me to visit a lab in Japan to learn a specific protocol that proved fundamental to my research. His passion for discoveries had no boundaries. It is important to remember that although we need role models and champions for women in science, we also need truly inspirational figures for science.

When my European grant ended, it was exclusively thanks to Tim that I got the extra funding guaranteeing me a good amount of time to continue my project after my maternity leave. At the time, we had several chats about my future. I had doubts about a career as a group leader, but Tim said I was good at

science and technically very skilful, giving me the encouragement that I needed.

After my career in research ended, I became an editor. I know the weight of words; however, facts and actions are much more important. A year after that first consortium conference, and wearing a different jumper, Tim and I again headed for the annual meeting. This time, I had a different kind of shock. Tim suggested that I should be the speaker, not him, because it was my research. Well, it was quite stressful. The members of that consortium included some of the most important cell-cycle scientists in Europe, but it was a great opportunity, and not only for me. Tim's action led to other group leaders leaving the stage to their postdoctoral fellows. Some were men and some were women. I doubt Tim cared which; all he saw were promising young scientists. ■

**Alessia Errico** is a former editor for Nature Reviews Clinical Oncology.  
e-mail: [alessia.errico@googlemail.com](mailto:alessia.errico@googlemail.com)

I KNOW THE  
**WEIGHT OF  
WORDS; HOWEVER,  
FACTS AND  
ACTIONS ARE  
MUCH MORE  
IMPORTANT.**

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/zsp1fr](http://go.nature.com/zsp1fr)



# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## ENERGY

### Rising cost of climate inaction

Increasing temperatures could boost electricity demand and costs over the next several decades, if warming continues unabated.

James McFarland at the US Environmental Protection Agency in Washington DC and his team analysed rising temperature trends and climate policies from 2015 to 2050 using three electricity-sector models for the United States. They found that electricity demand in 2050 would rise by 1.6–6.5% more in scenarios with increasing temperatures than in the case of constant temperatures, owing to factors such as greater use of air conditioning.

Without global action, the cost of electricity generation would grow by 1.7–8.3%, which is comparable to the projected costs of cutting carbon emissions from the power sector by 50% by 2050. *Climatic Change* <http://doi.org/5gh> (2015)

## CANCER

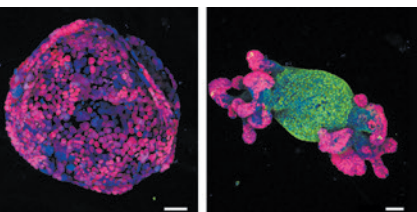
### Colon-cancer cells made normal

By turning on a particular gene, researchers have made colon-cancer cells in mice revert back to normal ones.

Scott Lowe at the Memorial Sloan Kettering Cancer Center in New York and his colleagues engineered mice so that they could use a small RNA

molecule to switch on the *Apc* gene, a tumour suppressor that is mutated in most colorectal cancers. Turning on this gene shrank precancerous polyps in the colon, and cells in the polyps developed and behaved normally. Even reactivating *Apc* in full-blown tumours bearing other cancer-driving mutations made the cells non-cancerous.

Silencing *Apc* in 3D cultures of mouse colon cells led to cancer-like growth (**pictured, left**; scale bars are 50  $\mu\text{m}$ ), whereas switching it back on restored normal cell division (**pictured, right**). Targeting the cell-signalling pathway



## CONSERVATION

### Elephant-poaching hotspots pinpointed

DNA analysis of seized ivory suggests that elephants have been poached at high rates in just two regions in Africa.

Samuel Wasser at the University of Washington in Seattle and his team studied genetic material from 28 large ivory seizures between 1996 and 2014 to identify the origins of poached tusks. They found that 86–93% of savannah-elephant tusks seized by authorities

since 2006 came from southeastern Tanzania and neighbouring northern Mozambique. A similar proportion of forest-elephant ivory originated from a region in central Africa: the nexus of Gabon, the Republic of Congo and the Central African Republic. The data could help authorities to focus regulatory and law-enforcement efforts on these regions, the authors say.

*Science* <http://doi.org/5h6> (2015)

## MICROSCOPY

### Graphene protects cells for imaging

A single-atom-thick mesh of carbon can protect living animal cells from being damaged under an electron microscope, and could lead to better cell images.

Tissue samples are typically dried and chemically treated to protect them from the vacuum of electron microscopes,

but this kills cells and can introduce structural artefacts. As an alternative, Ke Xu of the University of California, Berkeley, and his colleagues used graphene — which is conductive but impermeable to gas and liquid — to insulate live, untreated mammalian cells against the microscope's harsh environment. The team obtained images that correlated well with other live-cell imaging methods.

In another study, David Weitz of Harvard University in Cambridge, Massachusetts, and his team used 3–10 layers of graphene sheets to encapsulate liquid samples

ART WOLFE/WWW.ARTWOLFE.COM

KEVIN P. O'ROURKE

for transmission electron microscopy. They imaged influenza viruses and cells in their native state.  
**Nature Commun.** 6, 7384 (2015); **Nano Lett.** <http://doi.org/5gpg> (2015)

## CHEMISTRY

## X-rays make molecular movie

Researchers have made a film of a molecule's structural changes during a chemical reaction.

A team led by Michael Miniti at the SLAC National Accelerator Laboratory in Menlo Park, California, used a powerful free-electron laser to fire ultrafast X-ray pulses at the ring structure of 1,3-cyclohexadiene as it opened up to form the linear molecule 1,3,5-hexatriene. The scattered X-rays provided structural snapshots roughly every 80 femtoseconds during the 200-femtosecond reaction (1 femtosecond is  $10^{-15}$  seconds). Using these results and theoretical calculations, the team worked out the most likely path for the reaction.

The technique could reveal new reaction mechanisms, the authors say.

**Phys. Rev. Lett.** <http://dx.doi.org/10.1103/physrevlett.114.255501> (2015)

## ANIMAL BEHAVIOUR

## Kangaroos are lefties

Kangaroos that use two legs to jump have a strong preference for which hand they use to scratch themselves, suggesting that pronounced handedness is not uniquely human.

Yegor Malashichev of Saint Petersburg State University, Russia, and his team observed seven species of marsupial (including kangaroos) in Australia — three that walk on four legs and four species that are bipedal. They found that two kangaroo species

mostly used their left forelimbs for actions such as grooming (**pictured** is an Eastern grey kangaroo, *Macropus (Macropus) giganteus*). Red-necked wallabies (*Macropus (Notamacropus) rufogriseus*), which move on two legs, generally used their left forelimb for fine manipulation tasks, such as feeding, and the right forelimb for actions requiring greater physical strength, such as holding up branches when feeding.

Four-legged animals did not exhibit handedness, suggesting that two-legged locomotion is a prerequisite for this characteristic.

**Curr. Biol.** <http://doi.org/5jgh> (2015)

## PLANETARY SCIENCE

## Polar winds blow on Titan

Electrically charged particles stream away from Saturn's moon Titan, escaping into space in a similar way to Earth's polar atmosphere.

Titan's thick hydrocarbon haze is unique in the Solar System. Andrew Coates of University College London and his colleagues used the Cassini spacecraft to detect charged particles in Titan's uppermost atmosphere and deep space. Particles flowed away from the moon, along magnetic field lines that extended more than 17,000 kilometres from Titan.

The work explains how 7 tonnes of atmospheric molecules manage to escape from Titan every day.

**Geophys. Res. Lett.** <http://doi.org/5jg> (2015)

## GEOPHYSICS

## An island's ups and downs

An island off the coast of Chile lurched upward during earthquakes in 1835 and 2010, but subsided in between.

Such events provide a rare look at how Earth's crustal plates respond throughout an entire earthquake cycle.

## SOCIAL SELECTION

Popular topics on social media

## Curiosity over computer fact-checker

A computational fact-checker that can sort truth from fiction caught the attention of the online science community this week. Researchers at Indiana University in Bloomington mined Wikipedia's information boxes, which summarize key facts, to create a 'knowledge graph' of 3 million people, places and things. The resulting algorithm then used that knowledge to gauge the truth of statements that were presented to it, such as "Rome is the capital of Italy", with nearly the same accuracy as humans. The researchers acknowledge that the source material is not 100% reliable, as online commenters also noted. William Gunn, head of academic outreach for the research-management tool Mendeley in California, tweeted that the results might be different if the algorithm had been built using a different body of knowledge: "Wonder what this would look like using sci pubs as a knowledge graph?"

➔ **NATURE.COM**

For more on popular papers: [go.nature.com/tgjbym](http://go.nature.com/tgjbym)

Informatics researcher and lead author Giovanni Ciampaglia says that the team plans to use other sources of information to improve the fact-checker.  
**PLoS ONE** 10, e0128193 (2015)

Just after the 1835 quake, Robert FitzRoy of the HMS *Beagle* noted that seaweed and mussels had lifted as much as 3 metres above the shoreline of Santa María island. Now, a team led by Robert Wesson of the US Geological Survey in Denver, Colorado, has collected data from other old surveys and modern satellite measurements of the same island.

The scientists found that the island's height above sea level dropped by about 1.4 metres between 1835 and 2010, before rising 1.8 metres during the later quake. This suggests that strain builds up erratically where the Nazca and South American tectonic plates collide, indicating that the seismic activity in that region is more complicated than scientists had thought.

**Nature Geosci.** <http://dx.doi.org/10.1038/ngeo2468> (2015)

## CLIMATE-CHANGE ECOLOGY

## Plant diversity declines in the dry

The number of plant species in a California grassland area has dropped since 2000 as the area has become more arid — an

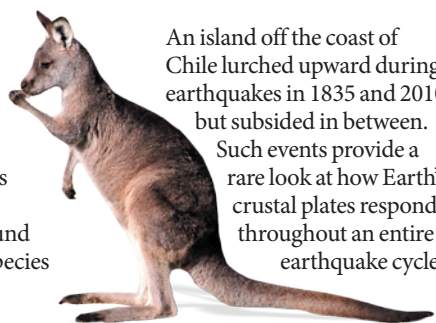
indication of how such ecosystems might respond to climate change.

Susan Harrison and her colleagues at the University of California, Davis, monitored plant diversity at 80 sites across a 27-square-kilometre area from 2000 to 2014. They found that the number of species at the sites and across the study area decreased over time, particularly in the number of native annual flowering plants. The decline in diversity correlated with diminishing precipitation during winter, when the annual plants are seedlings. The team did not find links between the fall in diversity and other possible causes such as fire, invasive species or nitrogen levels.

A reduction in precipitation is predicted for this area as the climate warms, so such biodiversity losses could signify future extinctions.  
**Proc. Natl Acad. Sci. USA** <http://dx.doi.org/10.1073/pnas.1502074112> (2015)

➔ **NATURE.COM**

For the latest research published by Nature visit: [www.nature.com/latestresearch](http://www.nature.com/latestresearch)





# SEVEN DAYS

The news in brief

## POLICY

### Japan whaling plan

The scientific committee of the International Whaling Commission (IWC) has been unable to reach a consensus on the merits of a proposed plan by Japan to resume hunting whales around Antarctica. Japan is seeking IWC permission for a new programme of whaling that it claims is for scientific purposes, but which critics claim is a commercial hunt in disguise. In a report published on 19 June, the committee said that it had not been given enough information to say whether killing whales was necessary to achieve Japan's stated scientific objectives. A previous hunt was deemed unscientific by an international court last year.

## RESEARCH

### Ebola trial ends

A clinical trial of an experimental Ebola drug being tested in people in Sierra Leone has stopped enrolment because the drug does not seem to be effective, the company leading the trial said on 19 June. The drug, TKM-Ebola-Guinea, consists of RNA molecules designed to stop the virus from replicating. Despite encouraging results in monkeys, the drug was deemed unlikely to show

## NUMBER CRUNCH

# \$95m

The amount that drug company GlaxoSmithKline will give over five years to set up a non-profit institute in Seattle, Washington, to study cell 'operating systems' and their role in health.



MIKE HILL/GETTY

## High seas to get legal protection

Biodiversity in ocean areas outside national jurisdictions — the 'high seas' — will receive protection under new laws, the United Nations said on 19 June. A resolution adopted by the UN General Assembly states that rules will be designed to ensure the "conservation and

sustainable use" of marine life in such regions, which are currently often unmanaged. Amid increasing scrutiny of fishing and other activities, a treaty that governs responsible use of the high seas has long been a goal of conservation organizations and some governments.

benefits in humans, said Tekmira, its manufacturer, based in Vancouver, Canada.

## PEOPLE

### Sandia head

Materials scientist and engineer Jill Hruby will become the first woman to head any of the three US national-security laboratories when she takes over the running of Sandia National Laboratories on 17 July. Her appointment was announced on 22 June. Hruby has been at Sandia, the main offices of which are in Albuquerque, New Mexico, and Livermore, California, for more than three decades, and has had roles overseeing nuclear, biological and chemical security. Like the Los Alamos and Lawrence Livermore

national laboratories, Sandia is trying to reinvent itself by focusing on scientific research for today's national security needs.

### French minister

Thierry Mandon was named French minister of state for higher education and research on 17 June, after the post had lain vacant for three months. He replaces Geneviève Fioraso, who stepped down in March for health reasons. Mandon is a seasoned politician, but from 1998 to 2014 he served as president of Genopole, a cluster of biotechnology and genetics companies and academic research labs in Evry, south of Paris. French researchers have welcomed Mandon's appointment. See [go.nature.com/yizy3d](http://go.nature.com/yizy3d) for more.

## FACILITIES

### Telescope restarts

Construction of the Thirty Meter Telescope (TMT) on the summit of Mauna Kea in Hawaii is to resume on 24 June. Work has been halted since early April because of protests by Native Hawaiian groups, who see it as a violation of a sacred mountain. "We are now comfortable that we can be better stewards and better neighbors during our temporary and limited use of this precious land," Henry Yang, chair of the TMT International Observatory Board, said in a 20 June statement. The controversy has caused Hawaii's governor to limit future astronomical development on Mauna Kea, which currently hosts 13 telescopes.

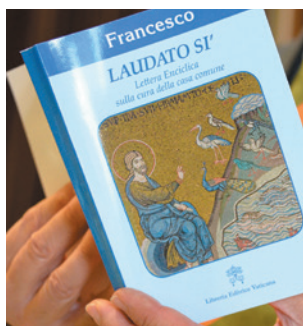
## EVENTS

## Imaging satellite

Sentinel-2A, an Earth-observation satellite that scientists say could revolutionize monitoring of land-use changes, launched on 22 June from Europe's Spaceport in Kourou, French Guiana. The satellite has superior specifications to Landsat-8, the most advanced US Earth-observation satellite so far, and will be joined in orbit by its twin next year. Together they will image the entire planet every two to five days. The duo is part of a planned fleet of six families of Sentinel satellites that will form the core of the European Union's Copernicus environmental monitoring system (see [go.nature.com/evfinv](http://go.nature.com/evfinv)), which is designed to provide near-real-time data on many environmental variables.

## Papal climate call

In a long-anticipated encyclical letter (**pictured**), Pope Francis warns that climate change caused mainly by human activity threatens to devastate ecosystems and human civilization. Released on 18 June, the nearly 200-page letter contains a sweeping account of scientific findings and urges world leaders to transition the global energy system away from fossil fuels.



Scientists and policy-makers hope that the encyclical will help to strengthen public trust in climate-change science among the more than 1 billion Catholics worldwide, and facilitate international efforts to cut global emissions of greenhouse gases. See [go.nature.com/r1i3q4](http://go.nature.com/r1i3q4) for more.

## Thai MERS case

Thailand announced its first case of Middle East respiratory syndrome (MERS) on 18 June. An elderly man, who travelled from Oman to Thailand and sought medical care, was diagnosed with MERS at a Thai hospital. The man and his family were put into isolation, and infection-control measures implemented. Meanwhile, the MERS outbreak in South Korea seems to be slowing down. As of 23 June, there have been 175 confirmed cases in South Korea and 1 in China. The South Korean Ministry of Health and Welfare announced on 21 June that

seven experts from the US Centers for Disease Control and Prevention would be visiting the country to provide technical assistance.

## End of Aquarius

A US–Argentinian research satellite that carried Aquarius, the first instrument specialized to measure ocean salinity from space, has stopped working, NASA announced on 17 June. A piece of hardware aboard the SAC-D observatory failed and sent the satellite spinning out of control earlier this month. Although Aquarius had finished its main three-year mission, it had been expected to last for at least another year. Among other achievements, Aquarius's maps show how fresh and salt water interact as rivers pour into the ocean.

## BUSINESS

## Hepatitis C drug

China's patent office has rejected attempts to patent a blockbuster hepatitis C drug, said charity Médecins Sans Frontières of Geneva, Switzerland, on 19 June. The antiviral drug sofosbuvir is made by Gilead Sciences of Foster City, California, but health advocates have argued that it is too expensive for people in middle- and low-income countries. It costs US\$84,000 per course, although Gilead has negotiated

## COMING UP

28 JUNE – 3 JULY

More than 650 young researchers meet 65 Nobel laureates for this year's Lindau Nobel Laureate Meeting in Germany. The scientists come from all three natural-science Nobel Prize disciplines: medicine and physiology, physics and chemistry. [go.nature.com/tpithk](http://go.nature.com/tpithk)

30 JUNE

The world's first Asteroid Day is held on the anniversary of the 1908 Tunguska event in Siberia, the largest impact on Earth in recent history. The event is intended to raise global awareness about the risk from asteroids and how to protect Earth. [go.nature.com/uvhz7w](http://go.nature.com/uvhz7w)

a discounted price of \$900 with countries including Egypt. China's decision — which has not been announced by the government — could spur other countries to challenge the patent, leading to the manufacture of cheaper generic versions.

## Vaccine purchase

Pharmaceutical giant GlaxoSmithKline (GSK) of Brentford, UK, has sold two meningitis vaccines to Pfizer in New York City for £82 million (US\$130 million). Announced on 22 June, the deal will give Pfizer Nimenrix and Mencevax, which protect against 4 strains of meningitis. The sale was deemed necessary by regulators after GSK's purchase earlier this year of a vaccines business owned by Novartis of Basel, Switzerland; that transaction included two other meningitis vaccines.

► [NATURE.COM](http://NATURE.COM)

For daily news updates see: [www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

The global market for photovoltaic solar energy will rise rapidly, says the consultancy GTM Research in Boston, Massachusetts. It predicts that by the end of 2015, the market will have expanded by 36% with 55 gigawatts (GW) of new installed solar capacity; southeast Asia will account for 30 GW of this. Total solar photovoltaic capacity will triple by 2020 to nearly 700 GW. By then, demand for solar panels will be largely independent of government incentives, says the analysis.

## SOLAR CAPACITY SOARS

The world's solar photovoltaic capacity is predicted to triple by 2020.





# NEWS IN FOCUS

**CONSERVATION** Treatments for deadly bat fungus could buy time **p.400**

**ASTEROID WATCH** Space-rock hunters struggle to raise cash **p.402**

**PUBLISHING ETHICS** Biomedicine offers lessons for Earth science **p.403**

**BIOMEDICINE** Paul Nurse gambles on giant new lab **p.406**



CARLO ALLEGRI/REUTERS/CORBIS



Trains are full of more than just people: bacteria stick to seats more effectively than they do to metal poles.

## MICROBIOLOGY

# Urban microbes unveiled

*Genomics tracks the microscopic inhabitants of homes, transport systems and sewers.*

BY RACHEL EHRENBERG

Embedded in the filth and chaos of the world's great metropolises, amid the people, pigeons, cockroaches and rats, there is a teeming world of bacteria, viruses, fungi and protists that scientists are only now surveying. Microbes are everywhere: on trains, pavements and lifts; in parks, libraries, hospitals and schools. Most are innocuous, some are friendly, and a handful cause death and disease. But the vast majority are unknown.

Researchers described results of early forays into this terra incognita at the Microbes in the City conference on 19 June, hosted by the New York Academy of Sciences and New York University (NYU) on the 40th floor of an antiseptic-looking glass office tower in Manhattan. "We're really at the infancy of a very interesting

scientific endeavour," said Joel Ackelsberg, a medical epidemiologist for the New York City Department of Health and Mental Hygiene. "Right now, we know very little."

Researchers are not even sure how to survey this strange landscape. There are competing techniques for detecting, quantifying and keeping track of which microbes are doing what in the built environment, and where. But researchers believe that efforts could lead to new approaches for monitoring bioterrorism, tracking disease outbreaks or assessing the impact of storms and pollution.

Each month, high-throughput techniques allow scientists to sequence roughly 1,000 microbial genomes from samples collected in various environments, said computational biologist Curtis Huttenhower of the Harvard T. H. Chan School of Public Health in Boston,

Massachusetts. That is an impressive amount of data, but it is dwarfed by the unfamiliar. Christopher Mason, a computational geneticist at Weill Cornell Medical College in New York City, told the conference how a baseline survey of genetic material from surfaces in the city's subway system had uncovered DNA from almost 1,700 known taxa, mostly harmless bacteria. But 48% of the genetic material did not match anything yet identified. "Half the world under our fingertips is unknown," said Mason.

Still, trends are emerging from the global Metagenomics and Metadesign of Subways and Urban Biomes initiative (MetaSUB), which aims to characterize the genetic material found on public-transport systems in 16 world cities to elucidate the microscopic riders that share the commute. Storms leave a mark: months after New York City's South Ferry Station was ►

► flooded in 2012's Hurricane Sandy, it still harboured DNA from bacteria associated with cold marine environments and fish, Mason said. However, most of the bacteria in the subway were harmless *Acinetobacter* species and others associated with human skin.

In his talk, Huttenhower described a survey of Boston's transit system that yielded similar flora. "Everything is covered in skin," he said. He noted that metal poles on the trains, which riders commonly consider hygienically suspect, actually retain much less bacterial biomass than the system's upholstered seats or plastic hand grips.

Microbiomes in houses tend to match those of the homes' human inhabitants — and quickly morph after a change in occupancy, said environmental microbiologist Jack Gilbert of Argonne National Laboratory in Illinois. He and his colleagues described results from a survey of ten homes, which found that they become populated with new residents' microbes within 24 hours.

Rodents are under study, too. White-footed mice (*Peromyscus leucopus*) in New York City carry more *Helicobacter* and *Atopobium* bacteria — associated with stomach ulcers and bacterial vaginosis in humans — than their suburban counterparts, but are totally free of tick-borne pathogens, reported biologist Alyssa Ammazalorso of the Albert Einstein College of Medicine in New York City. The city's rats carry a number of bacteria known to cause problems in people, said epidemiologist Ian Lipkin, director of the Center for Infection and Immunity at Columbia University in New York. He and others have found pathogenic *Escherichia coli*, *Clostridium difficile*, *Salmonella enterica* and the Seoul strain of hantavirus, which can be fatal when transmitted to humans (C. Firth *et al. mBio* 5, e01933-14; 2014).

Sewage samples from New York City's 14 wastewater-treatment plants turned up a disturbing number of genes for resistance to antibiotics, reported genomicists Susan Joseph and Jane Carlton of NYU's Center for Genomics and Systems Biology. As a rich human-derived soup spiked with antibiotics, sewage provides an ideal niche for the growth and spread of resistance, Joseph said. Martin Blaser, director of the Human Microbiome Program at the NYU School of Medicine, said that as populations of resistant microbes and their defensive tools become more diverse, the diversity of human-associated microbes in general is declining. He told how he and his colleagues have found that people in the West carry fewer protective bacteria than isolated human groups such as the Yanomami of the Amazon rainforest.

"We may have lost as much as half of our diversity already," said Blaser, "just as we are beginning to realize how important it might be." ■



The fungus that causes white-nose syndrome eats into bats' wings.

#### ECOLOGY

# Bacteria may help bats to fight fungus

*As white-nose syndrome spreads, researchers are trialling ways to stop colonies from collapsing.*

BY NALA ROGERS

The bats at Marm Kilpatrick's two Illinois field sites perished right on schedule. The mines sheltered nearly 30,000 bats before white-nose syndrome, a deadly fungal disease, arrived in late 2012. By March 2015, less than 5% remained.

Kilpatrick, a disease ecologist at the University of California, Santa Cruz (UCSC), and his colleagues chose the mines because they lay right in the path of the fungus, which has spread from Europe through 26 US states and 5 Canadian provinces since January 2007.

Although researchers are currently helpless to halt the spread of the fungus, there is reason for cautious optimism: treatments could soon be available that will help the bats to keep the infection at bay, for a season at least.

The goal, says Chris Cornelison, a microbiologist at Georgia State University in Atlanta, is to ensure that when researchers find long-term solutions for the disease, "there are still bats to treat".

The fungus (*Pseudogymnoascus destructans*) grows on bats while they hibernate in winter, digging into their noses, ears and wings. Animals that survive until spring usually clear the infection as their bodies warm; some species do it year after year. But the pathogen causes other species to repeatedly rouse from hibernation, which burns up fat reserves and can

cause the animals to starve to death. Some even flee their roosts in a futile search for food.

"They come out of caves in the winter, and they try to get into people's homes or churches or schools," says Jeremy Coleman, national coordinator for white-nose-syndrome research at the US Fish and Wildlife Service in Hadley, Massachusetts. "They're dead and dying on the ground."

Kilpatrick and his colleagues have discovered that a bacterium found on bats' wings may help them to combat infection. In April, the scientists published a paper in *PLoS ONE* showing that two strains of the bacterium *Pseudomonas fluorescens* kill white-nose fungus in cell culture (J. R. Hoyt *et al. PLoS ONE* 10, e0121329; 2015). Last winter, the researchers applied the treatment to bats in the lab. They have not released their results, but Kilpatrick hopes to test the bacteria in the wild soon.

Others are examining whether volatile organic compounds produced by *Rhodococcus* bacteria, which are found in soils, can kill white-nose spores. During a field trial in Missouri last winter, Cornelison and his colleagues treated bats with such compounds for 48 hours before returning the caged animals to their cave for 4 months to finish hibernating. On 19 May, the researchers released the bats that were free of disease — prompting media coverage of a potential 'cure'.

Cornelison cautions against such

US FWS/RYAN VON LINDEN, NY DEPT ENVIRONMENTAL CONSERVATION/SPL



celebration. His team is still analysing the data from the trial, and has not revealed how many bats were treated and released or how the controls fared. And even if this or the bacterial treatments are effective, they will be only short-term solutions. The fungus lingers on cave walls during the summer, and bats do not seem to develop immunity to it — so researchers would need to treat the animals every year to keep them from getting sick.

#### STOPGAP SOLUTIONS

Long-lasting solutions remain elusive. Some scientists hope to develop a vaccine, but have yet to work out how to trigger the animals' immune systems to destroy the pathogen, says Ken Field, an ecoimmunologist at Bucknell University in Lewisburg, Pennsylvania. Bats naturally produce antibodies to the fungus, but there is no evidence that these can help them to survive.

Other researchers are promoting a more radical long-term solution: altering airflow in mines where bats hibernate to make the sites less hospitable to the white-nose fungus. Places where bats survive infection tend to be relatively cool and dry. By opening new routes to the outside, researchers could cool and dehumidify the air in mines that are too warm and wet.

There is a chance that manipulating airflow could drive bats to abandon the habitat, says Kate Langwig, an ecologist at the UCSC. But she argues that it is worth a try, because at some sites the white-nose fungus kills about 90% of the bats present. In the first 5 years that it was present in the United States, the pathogen claimed more than 5.5 million animals.

In the meantime, the United States and Canada are developing and implementing strategies to coordinate work by scientists and by local and national governments — ranging from laboratory and field studies to efforts to prevent people from inadvertently spreading the fungus to pristine caves.

The plight of the bats is “stark — it’s demoralizing”, says Winifred Frick, an ecologist at the UCSC. “But I have hope in terms of the amount of creative energy and sense of dedication that people are putting forth on this problem. If there are solutions, we will find them.” ■

#### SPACE

# First glimpse of primordial stars

*Astronomers claim to spot generation that seeded Universe.*

BY ELIZABETH GIBNEY

Some of the first generation of stars, whose explosions breathed carbon, oxygen and other elements into the Universe, may have been glimpsed for the first time. The possibility comes as a pleasant surprise to astronomers, who did not expect to be able to spot these primordial objects with existing telescopes.

Primordial stars are theorized to be hundreds of times larger than the Sun, and made up only of pristine hydrogen, helium and traces of lithium left over from the Big Bang. The earliest specimens formed during the first few hundred million years of the Universe, living for only a few million years before exploding in supernovae that laid the seeds for the more element-rich stars to come. But they have never been seen.

A team led by David Sobral, an astronomer at the University of Lisbon, now reports that it may have spied a late-blooming cluster of such stars, in the brightest distant galaxy yet observed. The stars, seen as they were when the Universe was around 800 million years old (a mere 6% of its current age), seem to be primordial in composition — but, strangely, they also seem to reside in the same galaxy as some second-generation stars.

“Until now, work on these stars has been completely theoretical,” says Sobral. “For the first time, we’re starting to get observations that can test the many theories about these stars and begin to understand how they formed.” His team’s report has been posted on the preprint server arXiv and accepted for publication in *The Astrophysical Journal* (D. Sobral *et al.* Preprint at <http://arxiv.org/abs/1504.01734>; 2015).

Seeing these ancient stars involves observing very distant galaxies. Their light takes billions



Artist's impression of galaxy CR7.

of years to reach us, and so shows the Universe in its earliest days. But the light is faint, making it difficult to spot. The short lifetimes of the first stars also makes them hard to find.

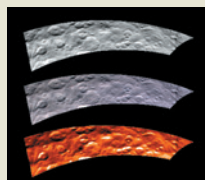
The surprise discovery emerged after Sobral and an international team of astronomers made a wide sweep of the sky using the Subaru Telescope on Mauna Kea in Hawaii. They used three further telescopes to peer into particularly bright galaxies, and found an intriguing signal from one that they named COSMOS Redshift 7. (The name was chosen so it could be abbreviated as CR7, to echo the nickname of Portuguese footballer Cristiano Ronaldo.)

The spectrum of light from CR7 showed evidence of ionized helium, which suggested that the source of the light was extremely hot. At such temperatures, any carbon and oxygen present should also have ionized, says Sobral. But there was no sign of these elements in ►



**MORE  
ONLINE**

#### TOP STORY



Intriguing geology of Ceres is revealed in new pictures  
[go.nature.com/bjzbva](http://go.nature.com/bjzbva)

#### MORE NEWS

- Nuclear monitoring agency reaches out to scientists [go.nature.com/bvtlme](http://go.nature.com/bvtlme)
- Europe's first humans: what scientists do and don't know [go.nature.com/usdonp](http://go.nature.com/usdonp)
- Pope or International Energy Agency — who said what? [go.nature.com/as3pr6](http://go.nature.com/as3pr6)

#### NATURE PODCAST



Antarctica's surprising biodiversity — and a new way to diagnose cancer early. [nature.com/nature/podcast](http://nature.com/nature/podcast)

► the light, which strongly suggested that it came from the first generation of stars.

CR7 also hosts second-generation stars, made from material spat out by dying first-generation stars, which means that it is not the sort of galaxy in which astronomers had imagined that they would make such a discovery. Sobral and his colleagues suggest that the primordial stars may be late developers, formed from a cloud of pristine and uncontaminated gas that was prevented from cooling and coalescing by the heat of strong radiation from earlier-blooming stars. This later formation

date would also explain why the stars were visible to Subaru.

That primordial stars should turn up in a large and evolved galaxy presents a challenge to the group's interpretation of CR7's light signal, but is probably the least exotic of the possible explanations, says Naoki Yoshida, an astrophysicist at the University of Tokyo. One alternative, that the emission comes from a supermassive black hole that formed directly from the collapse of a pristine cloud of gas, would actually be "even more spectacular", adds Sobral.

Until now, many astronomers believed that seeing the first generation of stars would take an instrument such as NASA's James Webb Space Telescope, which is projected to cost nearly US\$9 billion, is scheduled to launch in 2018 and should have the power to look further back in time than any previous instrument. But if Sobral and his colleagues are right, it may be possible to see primordial stars with existing telescopes. Indeed, similar bright galaxies have already been identified as potential candidates, adds Sobral. "We may already be looking at them." ■



An object estimated at 20 metres across exploded over Chelyabinsk, Russia, in 2013.

#### PLANETARY SCIENCE

# Private asteroid hunt lacks cash

*Foundation fails to raise funds it needs for a space telescope to catalogue near-Earth objects.*

BY TRACI WATSON

Astronomy and science organizations have declared 30 June as Asteroid Day, with plans to talk up the danger of asteroids that might be on a collision course with Earth. One partner in the effort, the Sentinel mission, has an especially urgent need to drum up public support: it is struggling to raise the US\$450 million it needs to launch a space telescope dedicated to finding hundreds of thousands of near-Earth objects.

Astronomers have long sought a spacecraft that would hunt for near-Earth objects

full-time, but public funding for such a programme has never materialized. Ground-based surveys have identified nearly all of the largest asteroids, but many of the space rocks that measure between 50 and a few hundred metres in diameter — big enough to wipe out a city with a direct hit — remain uncharted.

"Without space-based assets, progress is going to stall," says physicist Mark Boslough of Sandia National Laboratories in Albuquerque, New Mexico.

Sentinel, announced in 2012 by the B612 Foundation of Mill Valley, California, was supposed to provide that eye in the sky,

and NASA had hoped that the privately run effort would supplement the agency's own asteroid-hunting programme. NASA is considering funding an alternative mission, the Near-Earth Object Camera (NEOCam), initially proposed to the agency in 2006. But some scientists fear that Sentinel could spoil the chances of NASA supporting NEOCam. If Sentinel cannot fly, "then it's just a distraction from getting a job done that needs to get done," says planetary scientist Timothy Swindle of the University of Arizona in Tucson.

Ground-based surveys have located roughly 90% of the large asteroids that might be a concern — these are all bigger than one kilometre in diameter and pass by Earth's orbit at a distance of 45 million kilometres or less. An impact by one of these behemoths would have global consequences; fortunately, such immense objects plummet to Earth on average only once every 700,000 years or so, and none of those known is heading this way any time soon.

The statistics are less reassuring for smaller objects. As of August 2014, according to data in a recent paper, surveys had spotted only 565 near-Earth asteroids ranging from roughly 45 to 55 metres across — out of an estimated total of up to 520,000 (A. W. Harris and G. D'Abramo *Icarus* 257, 302–312; 2015). An asteroid no larger than that damaged trees across 2,000 square kilometres of Siberian forest in 1908, in what is known as the Tunguska event (see 'Direct hit'). Tunguska-sized bodies strike Earth once every 500 years or so, Boslough says. An object that exploded over Chelyabinsk, Russia, in February 2013 was estimated to have been about 20 metres in diameter.

#### PLANETARY PROTECTION

Sentinel aims to ensure that a similar fate does not unexpectedly befall a place the size of Lagos or London. Mission director Harold Reitsem, who helped to design instruments on the Hubble Space Telescope, says that Sentinel will use an infrared detector to find 90% of near-Earth objects larger than 140 metres across, and will also spot a good number of objects measuring 30 metres or more. If an asteroid were found to be headed for Earth long enough before impact, authorities could try to knock it off course.

AP PHOTO/AP VIDEO



Or they could plan to evacuate the target area.

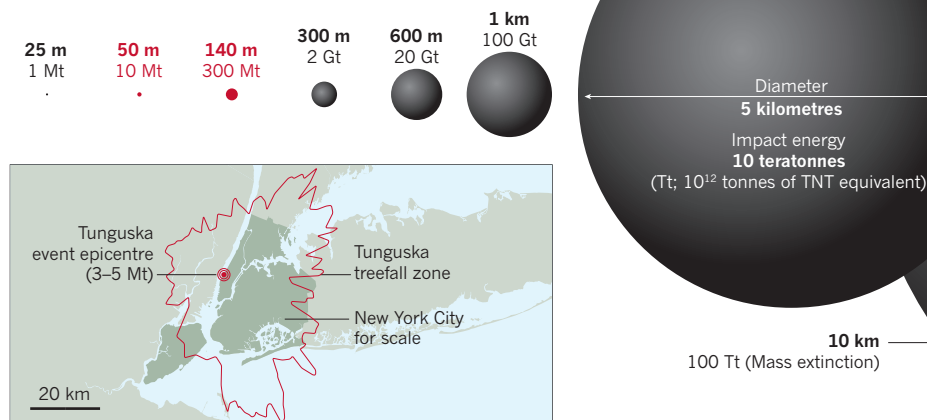
To help Sentinel along, in 2012 NASA signed an agreement with the B612 Foundation to provide analytical support and, once the spacecraft is in orbit, a data downlink. The space agency has an interest in Sentinel's success; it is under orders from Congress to find 90% of all near-Earth objects 140 metres or bigger by 2020. NASA will miss that target whether Sentinel launches or not, but a good space telescope could help to build a catalogue more quickly than ground-based surveys alone, which would need perhaps several decades.

"I don't want anybody to think that B612 is going to save the planet, but they are doing what we need to do," NASA administrator Charles Bolden said in 2013 when Congress pressed him on asteroid-detection efforts.

Yet progress has been slow. The B612 Foundation raised donations of roughly \$1.2 million in 2012 and \$1.6 million in 2013 — far short of its annual goal of \$30 million to \$40 million. NASA says that Sentinel has also missed every development milestone laid out in the 2012 agreement. In a January statement to an advisory panel, NASA said that its "reliance on the private sector for a space-based NEO survey ... is being re-examined". NASA's Lindley Johnson, director of the near-earth object programme, declined to speak to

## DIRECT HIT

In the 1908 Tunguska event, an object exploded over Siberia, damaging trees across 2,000 square kilometres. If a similar event occurred above New York City, the damage would hit all five boroughs and beyond. Impact energies for different sized objects are shown.



*Nature*, citing the ongoing discussions between the B612 Foundation and the agency.

NEOCam, meanwhile, would use an infrared telescope to search for asteroids from a vantage point between Earth and the Sun. In September, NASA will decide whether it is a finalist out of more than two dozen proposals being considered for launch by 2022 through the Discovery programme, which caps each

mission's cost at \$450 million.

If Sentinel receives substantial funding soon, it could launch by late 2019, says B612 Foundation chief executive and former astronaut Edward Lu. Even if NASA terminates its agreement with the foundation, he vows to keep the project going. "Believe me, I could do a lot of other things," he says. "But I feel this is extremely important." ■

## ETHICS

# Earth science wrestles with conflict-of-interest policies

*Industry-funding controversies highlight lack of standards among field's journals.*

BY JEFF TOLLEFSON

The undisclosed industry ties of some authors of Earth-science papers have raised ethical questions about how the field handles conflicts of interest. The cases of global-warming sceptic Willie Soon and hydrologist Donald Siegel have inspired calls for a uniform policy on reporting funding from entities that have an interest in the outcome of research.

"The Earth-science community doesn't really have a coherent set of policies for dealing with this," says Naomi Oreskes, a science historian at Harvard University in Cambridge, Massachusetts, who co-authored a 12 June commentary<sup>1</sup> in *Environmental Science & Technology* calling for stronger disclosure rules.

Soon, a solar physicist at the Harvard-Smithsonian Center for Astrophysics co-authored a 2010 paper<sup>2</sup> on climate-change policy in *Ecology Law Currents*, but he did not disclose funding from Southern Company, an electricity

provider in Atlanta, Georgia. The company has lobbied against stronger regulations to limit greenhouse-gas emissions.

The journal's editors told the Climate Investigations Center (CIC), a watchdog group in Alexandria, Virginia, that they do not have a conflict-of-interest policy, but are "exploring the possibility".

Soon also published a study<sup>3</sup> in March in *Nature Geoscience* on atmospheric conditions during the Little Ice Age — a 500-year cool period that began around the 1400s — without reporting his funding. Nature Publishing Group (which also publishes *Nature*) says that Soon complied with its policies, which require the disclosure of financial ties that are relevant to the research in question.

But even if the study had little relevance to climate policy, Oreskes says Soon's funders benefit from any paper in a peer-reviewed journal. Such funding relationships can create an unconscious bias and should be reported, she says.

Still, having disclosure policies does not guarantee that scientists will abide by them, as the CIC found when it examined other publications that Soon listed as "deliverables" in reports to Southern Company.

In 2009, Soon and others published a study<sup>4</sup> in the *Journal of Climate* on the variability of monsoons. The journal requires authors to disclose all funding sources and "any financial arrangement with a research sponsor that could give the appearance of a conflict of interest." Soon's co-authors acknowledged support from conventional granting agencies. But Soon did not report his funding from Southern Company. After the CIC released documents in February revealing Soon's industry ties, the journal amended the study to clarify his funding.

In another case, hydrologist Donald Siegel of Syracuse University in New York came under fire for a study<sup>5</sup> that concluded that fracking — which uses pressurized fluids to shatter rock and release trapped natural gas ▶

► — had not contaminated groundwater in Pennsylvania. Siegel did not disclose that the Chesapeake Energy Corporation of Oklahoma City, Oklahoma, had paid him and provided the water samples that his team analysed. *Environmental Science & Technology*, which published the analysis, posted a correction in April after the media revealed Siegel's links to the company.

Siegel says that he thought that this relationship was obvious, because he was working on a summer contract with the company's consultants, some of whom were listed as co-authors. "I never really anticipated this, but perhaps I was naive," Siegel says. "Because of the public nature of some of this science, I think we probably need a much tighter rein on what disclosure is."

Disclosure demands are coming from both sides of the aisle: climate sceptics have objected to a *Nature Climate Change* study<sup>6</sup> that analysed some of the US Environmental Protection Agency's greenhouse-gas regulations, because it was written by researchers who have received grants from the agency — even though the grants are listed in a public database.

Some experts suggest that Earth scientists should look to the biomedical community for guidance. Many biomedical journals require authors to fill out a common disclosure form that publishers developed in 1978 and have continued to update. And universities and hospitals often require medical researchers to report each year on their financial arrangements with industry.

Eric Campbell, a sociologist at Harvard's Edmond J. Safra Center for Ethics, says that biomedical scientists, funding agencies and journals have better-established disclosure policies because their work often involves human participants, and because of the strong financial ties between academia and the pharmaceutical industry.

There are no plans yet for the Earth-science community to develop disclosure standards. But there may be nascent backing for such an effort. Véronique Kiermer, director of author and reviewer services at Nature Publishing Group, says that the publisher "would be supportive of discussions in the Earth-sciences community about the specific challenges of the field and a framework for the standards of transparency it requires."

Campbell says that such a group discussion would be useful. "You don't want the individual with the conflict making decisions about what they should do about it." ■

1. Oreskes, N. *et al. Environ. Sci. Technol.* <http://dx.doi.org/10.1021/acs.est.5b02726> (2015).
2. Soon, W. & Legates, D. R. *Ecol. Law Curr.* **37**, 1–9 (2010).
3. Yan, H. *et al. Nature Geosci.* **8**, 315–320 (2015).
4. Liu, J. *et al. J. Clim.* **22**, 2356–2371 (2009).
5. Siegel, D. I. *et al. Environ. Sci. Technol.* **49**, 4106–4112 (2015).
6. Driscoll, C. T. *et al. Nature Clim. Change* **5**, 535–540 (2015).



ELAINE THOMPSON/AP

A plastic cast of the skull of Kennewick Man.

## ANCESTRY

# Genome results rekindle legal row

*'Kennewick Man' sequencing shows Native American roots.*

BY EWEN CALLAWAY

The genome of a famous 8,500-year-old North American skeleton, known as Kennewick Man, shows that he is closely related to Native American tribes that have for decades been seeking to bury his bones. The finding seems likely to rekindle a legal dispute between the tribes and the researchers who want to keep studying the skeleton. Yet it comes at a time when many scientists — including those studying Kennewick Man — are trying to move past such controversies by inviting Native Americans to take part in their research.

"The controversy has been painful for lots of people; tribal members and scientists as well," says Dennis O'Rourke, a biological anthropologist at the University of Utah in Salt Lake City.

Soon after the skeleton's bones were unearthed in 1996, near the shores of the Columbia River near Kennewick, Washington, several local tribes demanded the return of Kennewick Man, whom they dubbed the Ancient One. The US Army Corps of Engineers — the federal agency that manages the land where the remains were found — sided with the tribes,

citing a 1990 law that mandates the return of Native American remains and artefacts to affiliated tribes: the Native American Graves Protection and Repatriation Act (NAGPRA).

Several archaeologists and anthropologists sued the US government to stop the return, arguing that Kennewick Man was too old to be connected to the tribes. In 2002, a federal judge ruled that NAGPRA did not apply because the US government had not established that the tribes had a cultural affiliation with Kennewick Man. The decision was affirmed by a US Appeals Court in 2004.

A team led by Eske Willerslev, a palaeogenomicist at the Natural History Museum of Denmark in Copenhagen, used cutting-edge methods to extract snippets of DNA from a flake of finger bone and then sequence the genome. The researchers then compared the genome to that of members of dozens of groups from across North and South America, including several members of the Confederated Tribes of the Colville Reservation, one of the five groups seeking Kennewick Man's reburial (M. Rasmussen *et al. Nature* <http://doi.org/5jb>; 2015). Members of the



other four groups — the Nez Perce, the Umatilla, the Yakama and the Wanapum — opted not to contribute DNA to the study.

The Colville members were more closely related than were many North and South American tribes, but some other tribes still shared close ancestry with the 8,500-year-old. “We can’t say that the Colville are the closest living descendants of Kennewick Man, because the reference panel is too small,” warns Willerslev. “But I think we can say that Colville is very closely related to Kennewick Man.”

Gail Celmer, an archaeologist at the US Army Corps of Engineers in Portland, Oregon, says that the Colville people are still eager to pursue repatriation. Her agency therefore plans to reconsider whether Kennewick Man falls under NAGPRA. “We expect challenges, so we’re going to have to be very careful about how we do our reviews,” she says.

Hank Greely, a legal scholar at Stanford University in California, notes that with a genome sequence analysed, “it’s in the best interest of the scientists, of the government, and of Native Americans to think seriously about giving Kennewick Man’s remains back to the tribes”.

“The whole point of the Kennewick Man case was to ensure that important discoveries like this had an opportunity to be made,” says Doug Owsley, a forensic anthropologist at the

Smithsonian National Museum of Natural History in Washington DC and one of the scientists who sued the US government to gain access to the remains. “If I had to do it again, I would.”

James Chatters, an anthropologist who excavated Kennewick Man and who has since set up the consulting firm Applied Paleoscience in Bothell, Washington, is torn. “As a person who worked directly with the skeletal remains, I’d like to see them in peace,” he says. “As a scientist, I would hate to see one of the most complete North American skeletal remains be put back into the ground for political reasons.”

### CLOSE TIES

Willerslev is sticking to the sidelines on the issue of repatriation. “It’s somebody else who needs to figure that one out,” he says. But he wants the tribes to be involved in his research; after his team determined that Kennewick Man was Native American, he informed the five tribes about the conclusion. Some of the members travelled to his Copenhagen lab to learn more about the research, donning full-body suits to visit the clean lab in which ancient DNA is extracted from remains. The lab did similar outreach with tribes in Montana after it sequenced the genome of the ‘Anzick’ boy (M. Rasmussen *et al. Nature* **506**, 225–229; 2014), helping to broker a deal to rebury those bones.

Many other researchers are taking a similar

approach. O’Rourke says that there is no one-size-fits-all strategy to working with native communities. He finds some of the North American Arctic groups he works with eager to contribute to his research; others are less so, and their opinions shift over time.

“We really have to change the top-down approach, where we come to people and say ‘these are our research questions and you should participate, because — SCIENCE,’” says Jennifer Raff, an anthropological geneticist at the University of Texas at Austin.

Just weeks before Kennewick Man’s remains were discovered, researchers working in Alaska discovered a 10,000-year-old human skeleton. They notified local tribes and quickly came to an agreement that allowed them to excavate and study the remains and keep the tribes involved in the research. “You don’t really hear so much about the good cases,” says Raff. ■

See [go.nature.com/cnizsi](http://go.nature.com/cnizsi) for a longer version.

### CORRECTION

The News story ‘US “export rules” threaten research’ (*Nature* **522**, 266–267; 2015) should have said that information developed through fundamental research — rather than all unclassified information — is considered to be in the public domain.



# SIR PAUL'S CATHEDRAL

*When the Francis Crick Institute opens in London this year, it will be Europe's largest biomedical research centre. Can director **Paul Nurse** make this gamble pay off for UK science?*

BY EWEN CALLAWAY

Fifteen years ago, Paul Nurse had what he now calls “a really stupid idea”. The Nobel-prizewinning geneticist proposed bringing two of London's most esteemed biomedical-research centres under one big tent. Literally. He suggested building a megalab inside the Millennium Dome, a 1-kilometre-round, 52-metre-high open-plan exhibition centre with a plastic and fibreglass roof on the Greenwich Peninsula in southeast London. “It was crazy, because it was not built as a research institute,” says Nurse. “It got nowhere.”

Actually, Nurse's idea did go somewhere: about 10 kilometres northwest, right in the heart of London. In November, he will open the doors to the Francis Crick Institute, a lab even more ambitious than he originally conceived. When the £700-million (US\$1.1-billion), 93,000-square-metre institute reaches full capacity by 2021, it will house some 1,600 scientists and support staff, making it Europe's single largest biomedical lab.

Vast expectations come with that size. Nurse and the UK government have advertised the Crick as a boon to British science — one that will generate groundbreaking discoveries, lure the world's most brilliant young researchers to the United Kingdom and boost the British

life-sciences industry. And it will do so while eschewing departments, permanent positions and any sort of scientific focus. “It is quite a different sort of beast,” says Nurse. “I sometimes think not everybody realizes what's been put together here.”

If it works, this eccentric experiment in scientific infrastructure could guide how other countries form their labs of the future. But some question whether the Crick will have been worth the risk. They warn that the centre could become a black hole that devours public and private research funding, while widening the gap between the haves and have-nots of UK science. “It is a big, big, big, enormous whale of an investment,” says James Wilsdon, a science-policy researcher at the University of Sussex in Brighton, UK. “It has to succeed. If it doesn't, what do we do?”

## BIG AMBITIONS

The Crick was born of necessity. By the late 1990s, two of the United Kingdom's most esteemed biomedical-research institutions needed to find new homes. The pre-Second World War building that housed the National Institute for Medical Research (NIMR), in Mill Hill, north

PHOTO ILLUSTRATION: SKY; PETER GUDELLA/  
SHUTTERSTOCK; CRICK: ANNIHEA LEWIS/NATURE



REBECCA RED/EVERETT

London, was showing its age. So were some lab facilities supported by the Imperial Cancer Research Fund, a charity later renamed Cancer Research UK (CRUK). Nurse was head of that charity when he proposed the Millennium Dome megalab. But rather than pursue that dream, he left England in 2003 to lead the Rockefeller University in New York City.

Nurse's crazy idea lingered nonetheless. An influential 2006 government review, the Cooksey report, concluded that the country was "at risk of failing to reap the full economic, health and social benefits that the UK's public investment in health research should generate", and called for increased collaboration between funders and greater innovation in biomedical research. So when then-NIMR head Keith Peters revived Nurse's idea of merging the two institutions, it gained traction. Peters settled on a plot of land just north of central London, between the British Library and St Pancras railway station. To pay for the new centre, he brought in nearby University College London and the Wellcome Trust biomedical charity; these were later joined by King's College London and Imperial College London, which contributed £40 million each to construction costs. The universities will also house researchers at the Crick, especially physical scientists and engineers, and give the institute access to their associated hospitals.

The Crick, which is named after the co-discoverer of the structure of DNA, is billed as a basic-research institute. However, its funders are eager to see its discoveries turn into treatments. Researchers studying everything from stem-cell development to influenza will have access to drug-screening robots, high-powered microscopes and up to 200,000 animals, mostly mice and zebrafish.

The institute will evolve in its early years. The 90 or so group leaders currently at the NIMR and CRUK's London Research Institute (LRI) will be its first occupants when they start moving in early next year. The rest of the roughly 120 group-leader slots will go to young researchers recruited to work at the Crick for a maximum of 12 years — a 6-year initial stint and a 6-year extension, subject to a positive review. Researchers from the NIMR and the LRI who already have permanent posts will keep them, but as they retire, the proportion of tenured group leaders will shrink to around one-third of group leaders. The regular staff turnover, says Nurse, will free the Crick to follow the latest research trends, be they stem cells or genome editing. "A constantly rejuvenating institute stops things from becoming too inward-looking and ossified."

## LAB AS EXPERIMENT

"Gentle anarchy" is how Nurse has described his vision for the Crick. Researchers there will not belong to departments or divisions, and their location in the super-lab will depend on the core facilities they tend to use, from drug-screening robots to cell sorters. In Nurse's vision, researchers will form grass-roots interest groups on topics such as the microbiome. "Most people think it's crazy — you have no sorts of structures," says Nurse. "I think that it's very exciting and invigorating because the structure comes bottom up."

Tom Cech, a Nobel-prizewinning biochemist at the University of Colorado Boulder and former president of the Howard Hughes Medical Institute (HHMI) in Chevy Chase, Maryland, says that science funders in other nations will watch the Crick closely. He has some experience with unusual ventures. During Cech's tenure, the HHMI opened the Janelia Farm Research Campus, a neuroscience institute near Ashburn, Virginia, that kept group sizes small, jettisoned tenure and emphasized interdisciplinary research. "If nobody does these experiments, the question remains hypothetical. Experiments like Janelia Farm, like the Crick, are important for society," says Cech, because they can inspire new ways of doing science.

The Crick's supporters are also counting on the institute to be an economic engine for Britain by bridging the gap between the lab and market. Nurse's number two, chief operating officer David Roblin, was head of European research and development at pharmaceutical giant Pfizer until 2011, and has since worked for smaller biotechnology companies. "Although I'm an industrial scientist, my instructions are not to make it a pharmaceutical company," Roblin says. "The Crick is a discovery research institute that is very interested in translation." Roblin



Paul Nurse inside the Francis Crick Institute.

**"MOST PEOPLE THINK IT'S CRAZY — YOU HAVE NO SORTS OF STRUCTURES. I THINK THAT IT'S VERY EXCITING."**

is trying to lure pharmaceutical companies to place researchers at the Crick in the hope of speeding the transition to the clinic. But unlike at a private company, any developments would be openly reported. And the intellectual property would belong to the Crick.

Even in a city filled with glass and steel megastructures, the Francis Crick Institute is an imposing landmark. Nurse, who was knighted in 1999, says that his employees sometimes call it "Sir Paul's cathedral".

Perched on the edge of the city's revitalized King's Cross area, the building dwarfs the neighbouring British Library — the biggest public structure erected in Britain in the twentieth century. The Crick's twin, curved roofs give the appearance of an overturned ship, its hull split in two.

The institute feels even larger inside, where hard hats, safety glasses and steel-toed boots were still mandatory as of May. Hallways of glass-fronted open-plan labs stretch almost the length of two football fields on either side of an expansive atrium, over several lofty floors. To orient people, the walls are coloured by floor and a smartphone app is being developed ("I think we're blue. I can't remember," says one scientist who will move her lab to the Crick next year). Its four basement levels will house sensitive equipment such as electron microscopes, shielded by concrete slabs to block the ever-present vibrations from nearby train and tube lines. Construction has begun on a high-containment lab for working on influenza and other deadly pathogens, as well as facilities for a menagerie of lab animals, including the only collection of opossums used for research in the United Kingdom.

Like many new laboratories, the Crick is designing spaces to encourage interactions between researchers. Each above-ground floor centres around meeting rooms, tea and coffee stations and 'collaboration spaces', meant to increase the odds of fruitful encounters.

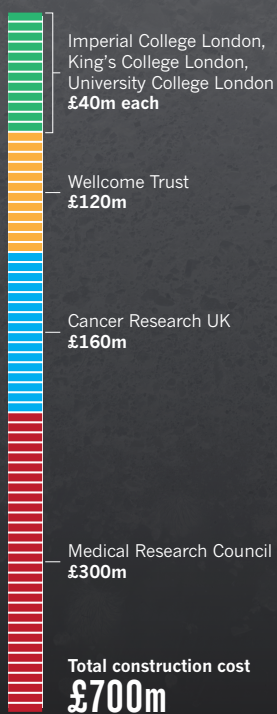
"That's one of the things I'm most excited about," says James Briscoe, a developmental biologist who is moving to the Crick from the NIMR. He has already met the group leaders with whom he will soon share bench space and equipment. Caetano Reis e Sousa, an immunologist coming from the LRI, says that he, like most other principal investigators,

## Vital signs

As Europe's largest biomedical centre, the Francis Crick Institute looms large in many dimensions, from the size of its budget to the number of its personnel.

### Money matters

On top of its £700-million (US\$1.1-billion) construction costs (below), the centre will have an annual budget of £150 million.



### Room to roam



already works with colleagues elsewhere. It is the graduate students, postdocs and technicians who will benefit most from the Crick's social engineering, he says. "They are the ones who will establish collaborations and connections that will allow new things to be done."

However, not everyone buys the argument that a giant fish bowl is needed for immunologists to have tea with physicists, the kind of interaction embodied by an unofficial mantra in many Crick documents and videos: "discovery without boundaries". The phrase grates on some researchers. "The assumption is, previously, discovery had terrible boundaries," says Kieron Flanagan, a science and technology policy researcher at the University of Manchester, UK. "And now that they're all in one building, suddenly the discoveries will flow. It's just bollocks."

The Crick's size is another sticking point. "Everybody believes that science has to get bigger and more capital intensive," says Flanagan. "There's a good bit of evidence that says small groups are more productive than large ones."

And its location presents challenges. Building the Crick in one of Europe's biggest cities will be crucial to convincing the best young researchers from around the world to move to the United Kingdom, says Nurse. The city's hospitals offer a population of potential clinical-trial participants not matched anywhere else in the United Kingdom.

But the institute has little room to expand, and has built several facilities off site, including a data centre and animal-breeding facility. A planned trans-London underground railway line, proposed to pass near the site, could overwhelm the vibration-dampening protections and jiggle the institute's sensitive scientific instruments (see *Nature* 518, 464–465; 2015). And then there is the cost of living in London, which has skyrocketed since plans for the Crick were unveiled. Nurse had hoped to buy a block of flats for postdocs and graduate students, but he could not find the money. He is now working on smaller

measures to make London more affordable for Crick scientists.

Gerald Rubin, executive director of the Janelia campus, sees the lack of subsidized housing as the institute's only weakness — but worries that it could hamper recruitment. "If the Crick had that, they would probably rapidly become the premier research institute in Europe. If they don't, I think they'll struggle," he says.

## RIPPLE EFFECTS

Nurse says that the Crick will serve as a scientific incubator for the United Kingdom: he hopes that researchers leaving the Crick will take their talents to universities throughout the country. "I say half joking that they should fall in love with somebody who lives here," he says.

Critics, however, say that Nurse is being glib about the serious economic challenges faced by young Londoners (a postdoc or even principal investigator's salary will not buy a flat anywhere near the institute). And they question whether the Crick will be the incubator that Nurse and the government contend. "I don't see it as being possible to define that as a strategy for the institute — to populate the UK," says David Stephens, a cell biologist at the University of Bristol. "It could, but there's no de facto reason why it will." Researchers could just as well move to world-class centres in the United States, Germany or Asia, say Stephens and others.

Some worry that the Crick will become 'too big to fail' — that its public and private funders will put the centre ahead of other priorities. Beyond its £700-million construction costs, the institute will have an annual budget topping £150 million per year (see 'Vital signs'), which will be paid by the government's Medical Research Council (MRC), CRUK and the Wellcome Trust. A flat or even diminished government science budget — increasingly likely after the election this year of a Conservative government that promises cuts — has amplified such concerns. "Taking a large scoop out of the pot for one institute makes it harder for the rest of us to attract funding," says Stephens, whose lab is supported by the MRC and the Wellcome Trust. Meanwhile, Nurse is currently leading a government review of how the UK research councils (which include the Crick's major funder, the MRC) divvy up money, raising conflict-of-interest concerns.

Nurse says that such complaints are unfounded "scare stories" from scientists who are "terrified about their own funding". He sees the scale of the Crick as a strength, which will help the institute in terms of both conducting world-leading research and securing financial support. "Having a single large institute, which has been invested in and where people have got a lot of skin in the game, is indeed a good encouragement to make the thing work," he says. "But it means we have a responsibility to deliver."

But deliver what? Gauging the ultimate success or failure of the Crick will be difficult. The researchers who will soon move into the institute are already leading scientists, and will continue to produce high-level work. But that is not enough. Cech says that the Crick should be judged on whether it contributes to discoveries that could not have been made by university researchers with the same funds. Wilsdon agrees: "It's going to have a lot of resources in it. It needs to do more than marginally outperform its competitors."

Nurse is another question mark. He plans to helm the Crick for another five or six years before stepping aside as director. Matthew Freeman, a molecular biologist at the University of Oxford, UK, wonders whether a successor will be able to carry his vision forward. "My sense is this is very much designed in a way Paul wanted it to be. He's put a lot of his own personality and weight and charisma behind establishing this rather unusual model," says Freeman. "I see this as an exciting experiment, but experiments can fail as well as succeed."

Nurse concedes that the Crick's loftier goals — creating a nimble, highly collaborative basic-research institute that powers Britain's knowledge economy — are less than certain. "I think we've got a reasonable chance," he says. But he is open to change. "What if my lack of departments is chaos? If it's chaos, I'll put the white flag up and say, right, we'll do something else," he says. "I am not a zealot." ■

Ewen Callaway writes for *Nature* from London.



# THE HARD SCIENCE OF OXYTOCIN

*As researchers work out how oxytocin affects the brain, the hormone is shedding its reputation as a simple cuddle chemical.*

In April 2011, Robert Froemke and his team were reprogramming the brains of virgin mice with a single hormone injection.

BY HELEN SHEN

a wave of more ambitious trials to test whether oxytocin can help some types of autism.

Before the treatment, the female mice were largely indifferent to the cries of a distressed baby, and were even known to trample over them. But after an injection of oxytocin, the mice started to respond more like mothers, picking up the mewling pup in their mouths. Froemke, a neuroscientist at New York University's Langone Medical Center in New York City, was monitoring the animals' brains to find out why that happened.

At first, the mice showed an irregular smattering of neural impulses when they heard the baby's cries. Then, as the oxytocin kicked in, the signal evolved into a more orderly pattern typical of a maternal brain. The study showed in unusual detail how the hormone changed the behaviour of neurons<sup>1</sup>. "Oxytocin is helping to transform the brain, to make it respond to those pup calls," Froemke says.

Oxytocin has been of keen interest to neuroscientists since the 1970s, when studies started to show that it could drive maternal behaviour and social attachment in various species. Its involvement in a range of social behaviours<sup>2</sup>, including monogamy in voles, mother-infant bonding in sheep, and even trust between humans, has earned it a reputation as the 'hug hormone'. "People just concluded it was a bonding molecule, a cuddling hormone, and that's the pervasive view in the popular press," says Larry Young, a neuroscientist at Emory University in Atlanta, Georgia, who has been studying the molecule since the 1990s.

That view has led some clinicians to try oxytocin as a treatment for psychiatric conditions such as autism spectrum disorder. But the early trials have had mixed results, and scientists are now seeking a deeper understanding of oxytocin and how it works in the brain. Researchers such as Froemke are showing that the hormone boosts neuronal signals in a way that could accentuate socially relevant input such as distress calls or possibly facial expressions. And clinical researchers are starting

The work is leading to a more sophisticated view of the hormone and its complex effects on behaviour — one that will take many types of expertise to refine. "The oxytocin field has just matured and ripened enough to draw in researchers from traditionally separate fields, catapulting this forward," says Young.

## BIRTH ACCELERATOR

Oxytocin's story starts back in the early 1900s, when biochemists discovered that a substance from the posterior pituitary gland could promote labour contractions and lactation. When scientists later discovered the hormone responsible, they named it oxytocin after the Greek phrase meaning 'rapid birth'. Oxytocin is produced mainly by the brain's hypothalamus; in the 1970s, studies revealed that oxytocin-producing neurons send signals throughout the brain, suggesting that the hormone had a role in regulating behaviour.

In a landmark 1979 study<sup>3</sup>, Cort Pedersen and Arthur Prange at the University of North Carolina in Chapel Hill showed that giving oxytocin to virgin rats could trigger maternal behaviours: the animals would build nests, lick or crouch over unfamiliar pups and even return lost pups to the nest. Researchers went on to show that oxytocin signalling in the brains of prairie voles (*Microtus ochrogaster*) helps the animals to form lifelong pair bonds<sup>4</sup> — a rarity among mammals. In 2012, researchers even found a version of oxytocin in the tiny roundworm *Caenorhabditis elegans*, where it helps the animals find and recognize mates<sup>5</sup>.

"This is a very ancient molecule," says Sue Carter, a neuroscientist at Indiana University in Bloomington, whose lab pioneered many of the early studies of oxytocin in voles. "It has been used and reused for many purposes across the evolution of modern animals, and almost everybody who's tried to look at an effect of oxytocin on anything like social behaviour has found something."

In mammals, many mysteries remain. Oxytocin is difficult to





measure reliably in the brain, making it hard to know exactly where, when and how much is normally released; nor do scientists understand precisely how it works to alter behaviour. “What we need to start thinking about is the more fundamental role that oxytocin plays in the brain,” Young says. The determination to find out has been strengthened by a growing move in neuroscience to characterize circuits that are important in brain operations. “That’s the level that’s critical for understanding how the brain is regulating behaviour,” says Thomas Insel, director of the US National Institute of Mental Health in Bethesda, Maryland, who has studied oxytocin in voles.

At Langone, Froemke focused on the circuits underlying the maternal response to pup cries — a behaviour that helps females to retrieve helpless newborns that can get lost when a mother is moving her nest. He focused on the left auditory cortex, a brain area thought to be involved in detecting the pups’ ultrasonic cries.

Froemke’s study<sup>1</sup>, published in April, showed that oxytocin temporarily suppresses inhibitory neurons — those that dampen neural activity — which allows excitatory cells to respond more strongly and reliably. “Our hypothesis is that the virgin brain is a blanket of inhibition, and that pairing the pup calls with oxytocin allows the network to be reconfigured,” says Froemke. The hormone may serve to amplify incoming signals and allow them to be recognized as behaviourally important. (It is at least possible, he says, that this same mechanism could explain why some human mothers feel they are uniquely tuned to a baby’s cries.)

“The study is kind of a high-water mark for the field, putting different levels all together: a robust behaviour, a brain region, and a cellular basis for it,” says Richard Tsien, a neuroscientist also at Langone. Tsien has been studying the action of oxytocin on neuronal circuits in detail, by examining slices of the hippocampus, a region involved in learning

**“WHAT WE NEED TO  
START THINKING  
ABOUT IS THE MORE  
FUNDAMENTAL ROLE  
THAT OXYTOCIN HAS  
IN THE BRAIN.”**

and memory. In a 2013 study<sup>6</sup> of rats, Tsien’s team found that oxytocin selectively acts on a type of cell called an inhibitory interneuron in a way that quiets background chatter within the neuronal circuit. “Oxytocin improved signal transmission, almost doubling the ability of information to flow through the system,” Tsien says. In effect, it is producing more signal and less noise.

Froemke’s and Tsien’s work fits into a broader theory: that one way oxytocin helps social interaction and recognition is by enhancing the brain’s response to socially relevant sights, sounds or other stimuli. Young has shown that the hormone helps mice to recognize and pay attention to the smells of other mice<sup>7</sup>; others found that it promotes people’s ability to recognize faces<sup>8</sup>.

The hormone does not act alone. In 2013, neuroscientist Robert Malenka at Stanford University in California and his colleagues showed that oxytocin works together with the neurotransmitter serotonin to reduce the excitability of neurons in the nucleus accumbens<sup>9</sup>, a brain region involved in reward. This process seems to support the preference of mice to return to environments where they had rewarding social interactions with other animals. “Oxytocin is part of a system,” Carter says, “and it’s not the only molecule that matters, but it’s one that in some way is regulatory over a large number of other systems.”

#### **MATTER OF TRUST**

The rapid evolution in basic research has been accompanied by a boom in clinical interest. Oxytocin has been used since the 1950s to accelerate childbirth, so many researchers consider it relatively safe to use in experiments.

About ten years ago, psychology studies started to show that single doses of oxytocin, delivered through an intranasal spray, could promote various aspects of social behaviour in healthy adults. People who inhaled oxytocin before playing an investment game were more willing



to entrust their money to a stranger than were placebo-treated players<sup>10</sup>. A dose of the hormone also increased the amount of time that people spent gazing at the eye region of faces<sup>11</sup>, and improved their ability to infer the emotional state of others from subtle expressions<sup>12</sup>.

The idea that oxytocin is central to social cognition made it an attractive candidate for treating psychiatric disorders, especially autism spectrum disorder. People with this condition, who often have problems with social interaction and communication, may not process social stimuli appropriately — and scientists theorized that oxytocin might reverse some of the symptoms. Beginning in 2010, results emerged that seemed to support this theory: researchers found that single puffs of oxytocin could temporarily improve measures of empathy and social cooperation in people with autism spectrum disorder.

“People got quite excited,” recalls clinical neuroscientist Evdokia Anagnostou, who co-directs the Autism Research Centre at Holland Bloorview Kids Rehabilitation Hospital in Toronto, Canada. But Anagnostou says that some preliminary steps were skipped over as researchers rushed to test oxytocin as a psychiatric drug. “To be honest, if we had done it properly, we wouldn’t have done it the way we did. It went a little bit too fast,” she says. Because oxytocin had cleared the early, standard steps of drug development decades earlier, some researchers did not systematically test a range of doses to see whether they had differing psychological effects.

Many early studies of oxytocin for autism were limited because they assessed only a single dose and had relatively few participants, and later experiments with more doses failed to show the same promise. In 2010, clinical psychologist Adam Guastella at the University of Sydney in Australia studied 16 male adolescents with autism spectrum disorder, and found that one dose of oxytocin could improve their ability to gauge the emotions of others by looking at their eyes<sup>13</sup>. But when he tried giving twice-daily doses of the hormone for two months, he found no significant improvements in social interaction or social cognition<sup>14</sup>. “Studies to this point have really shown limited benefit of oxytocin in improving psychiatric illnesses over time,” he says. Guastella says that getting to the bottom of oxytocin’s complex neurological effects will take time. “If we want a simple answer, we’re not going to get it.”

## ALL IN THE DETAIL

So far, few studies have definitively linked autism to problems in oxytocin signalling. Some of the clearest evidence emerged in February, from a team led by neurogeneticist Daniel Geschwind of the University of California, Los Angeles. The group showed that mice that lacked a working copy of the *Cntnap2* gene — which has been implicated in a small subset of human autism cases — had fewer oxytocin-containing neurons in the hypothalamus and socialized less with other mice than did control mice<sup>15</sup>. After receiving doses of oxytocin every day for two weeks, the mice behaved normally again. “Until this, there was no evidence that there was a subtype of autism that had to do with oxytocin deficits,” Geschwind says.

His study points to a more targeted approach in the clinic. “Autism is highly heterogeneous, but if you can find subsets of individuals — those who have oxytocin-signalling deficits — they may be the best candidates for oxytocin therapy,” says Karen Parker, a behavioural neuroscientist at Stanford.

A handful of large-scale clinical trials are now getting under way to test oxytocin and oxytocin-based therapies for autism spectrum disorder, and to work out who could benefit. Linmarie Sikich, a child psychiatrist at the University of North Carolina is heading the largest of these trials. Sikich plans to recruit 300 people with autism spectrum disorder, ranging in age from 3 to 17, and give them 6 months of either oxytocin or a placebo, followed by 6 months in which everyone will receive oxytocin.

Unlike previous studies, the trial will include people with a wide range

of symptoms — and one of its major aims is to uncover the set of factors that influence whether and how strongly people respond to oxytocin. Sikich will analyse many measures of cognition and social functioning, and collect blood samples to look for biomarkers — such as levels of oxytocin and the receptor it binds to — that are associated with a response. “Lin has really been trying to create conditions under which you could study the potential beneficial effects of oxytocin and really do this right,” says Carter.

But Carter and other scientists are concerned by reports from the physicians and parents of children with autism spectrum disorder who say that they are already using oxytocin off-label — before it has been thoroughly tested. “We do not understand how the hormone works yet, or have enough information about what happens when it’s given repeatedly,” Carter says. “This is not a molecule that people should be self-administering or playing with.”

Some work has pointed to a potential dark side to oxytocin. Carter’s group found that a single low dose of the hormone given to baby prairie voles improved their pair bonding as adults, but that higher doses

interfered with that behaviour — possibly because oxytocin started to activate other receptors<sup>16</sup>. And human studies have suggested that in certain contexts, a puff of oxytocin can cause people to be more aggressive in defending themselves against outsiders or competitors<sup>17</sup>. In patients with a psychiatric condition known as borderline personality disorder, a single dose of oxytocin has been found to hinder trust and cooperation<sup>18</sup>.

Young says that the oxytocin field would benefit from closer collaboration between basic and clinical researchers. If basic scientists can work out how oxytocin helps the brain to process social stimuli, then that might help in the design of stimuli — in the form of behavioural therapies — that could be given alongside the hormone to change behaviour, just as oxytocin and pup calls together affect virgin mice. “I think in the future these two branches need to have more communication,” Young says.

But long before that, say researchers, oxytocin could use a rebranding. “It doesn’t induce love; it doesn’t induce massive amounts of trust,” Guastella says. “The problem we’ve got ourselves into is that we’re trying to look for a simple answer: either oxytocin does or does not work in a patient population, or it does or does not enhance a certain social process.”

But the science of life is rarely as simple as that. “Oxytocin is known to affect circuits in different ways, and it’s not going to affect everyone in the same way,” Guastella says. “The sorts of biology we’re studying here are incredibly complex.” ■

**Helen Shen** is a science writer based in Sunnyvale, California.

1. Marlin, B. J., Mitre, M., D’amour, J. A., Chao, M. V. & Froemke, R. C. *Nature* **520**, 499–504 (2015).
2. Ross, H. E. & Young, L. J. *Front. Neuroendocrinol.* **30**, 534–547 (2009).
3. Pedersen, C. A. & Prange, A. J. *Jr. Proc. Natl Acad. Sci. USA* **76**, 6661–6665 (1979).
4. Williams, J. R., Carter, C. S. & Insel, T. *Ann. NY Acad. Sci.* **652**, 487–489 (1992).
5. Garrison, J. L. *et al. Science* **338**, 540–543 (2012).
6. Owen, S. F. *et al. Nature* **500**, 458–462 (2013).
7. Ferguson, J. N., Aldag, J. M., Insel, T. R. & Young, L. J. *J. Neurosci.* **21**, 8278–8285 (2001).
8. Rimmele, U., Hediger, K., Heinrichs, M. & Klaver, P. J. *J. Neurosci.* **29**, 38–42 (2009).
9. Dölen, G., Darvishzadeh, A., Huang, K. W. & Malenka, R. C. *Nature* **501**, 179–184 (2013).
10. Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U. & Fehr, E. *Nature* **435**, 673–676 (2005).
11. Guastella, A. J., Mitchell, P. B. & Dadds, M. R. *Biol. Psychiatry* **63**, 3–5 (2008).
12. Domes, G., Heinrichs, M., Michel, A., Berger, C. & Herpertz, S. C. *Biol. Psychiatry* **61**, 731–733 (2007).
13. Guastella, A. J. *et al. Biol. Psychiatry* **67**, 692–694 (2010).
14. Guastella, A. J. *et al. J. Child Psychol. Psychiatry* **56**, 444–452 (2015).
15. Peñagarikano, O. *et al. Sci. Transl. Med.* **7**, 271ra8 (2015).
16. Bales, K. L. *et al. Horm. Behav.* **52**, 274–279 (2007).
17. De Dreu, C. K. W. *et al. Science* **328**, 1408–1411 (2010).
18. Bartz, J. *et al. Soc. Cogn. Affect. Neurosci.* **6**, 556–563 (2011).

# COMMENT

**BIOETHICS** Crude gender politics debases the gene-editing debate **p.415**

**LAW** The struggle to save Argentina's glaciers from exploitation **p.416**

**FILM** Two films bring the asteroid threat to a new audience **p.418**



**OBITUARY** John Nash, game-theory Nobel prizewinner, remembered **p.420**

ILLUSTRATION BY DAVID PARKINS



## Science can't solve it

Democratically weighing up the benefits and risks of gene editing and artificial intelligence is a political endeavour, not an academic one, says **Daniel Sarewitz**.

**T**his year, several leading researchers have sounded warnings about the risks of using the CRISPR gene-editing technique to modify human<sup>1</sup> and other species' genomes in ways that could have "unpredictable effects on future generations"<sup>2</sup> and "profound implications for our relationship to nature" (see [go.nature.com/jq5sik](http://go.nature.com/jq5sik)).

Concerns are coming from the silicon sector as well. Last year, the physicist Stephen Hawking proclaimed that rapidly advancing artificial intelligence (AI) could destroy the human race. And in 2013, former Royal Society president Martin Rees co-founded the Centre for the Study of Existential Risk at the University of Cambridge, UK, in part to study threats from advanced AI.

Leaders of the scientific community are ready to share the responsibility for these powerful technologies with the public. George Church, a geneticist at Harvard University in Cambridge, Massachusetts,

and others wrote last year of CRISPR that "the decision of when and where to apply this technology, and for what purposes, will be in our collective hands".

But scientists also want to control the terms of engagement. The US National Academies, for example, will "guide decision making" by convening researchers and other experts later this year "to explore the scientific, ethical and policy issues associated with human gene-editing research". Scientists also emphasize the need for more research on risks and benefits to "better inform future public conversations"<sup>3</sup>. For instance, in the past few months, hundreds of scientists and technologists have signed an online open letter arguing that research

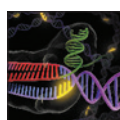
is necessary to learn how to accentuate the positive aspects of AI and avoid its potential perils (see [go.nature.com/jcyjib](http://go.nature.com/jcyjib)).

### WHO VALUES WHAT

The idea that the risks, benefits and ethical challenges of these emerging technologies are something to be decided by experts is wrong-headed, futile and self-defeating. It misunderstands the role of science in public discussions about technological risk. It seriously underestimates the democratic sources of science's vitality and the capacities of democratic deliberation. And it will further delegitimize and politicize science in modern societies.

The never-ending debates about genetically modified (GM) organisms, nuclear power, chemical toxicity and the efficacy of cancer screening should be evidence enough that science does not limit or resolve controversies about risk.

There is no way to capture the full ▶



**CRISPR GENE EDITING**  
A Nature collection  
[nature.com/crispr](http://nature.com/crispr)



► complexity of these issues from a scientific perspective. When new technologies are introduced into complex socio-technical systems, everyone is ill-informed about the risks. Differing bodies of evidence provide ammunition for competing views. Legitimate experts are always available to support conflicting preferences.

For example, an agricultural economist (concerned about crop yield) and an ecologist (concerned about ecosystems), will bring different sets of evidence, and probably entirely different values, to bear on studying the impacts of GM organisms. Even among agricultural economists, some researchers prefer field trials that allow for the careful control of variables such as weather and soil type; others study actual farms to capture real-world variability. These two perspectives often yield contradictory results<sup>4</sup>.

To many European consumers, moreover, research on crop yields is irrelevant. They are concerned<sup>5</sup> with the motives behind corporate decisions about crop varieties, aesthetic qualities of landscapes and food varieties, and principles of choice and transparency that would demand the labelling of GM foods even if there is no known health risk.

In other words, risk is more a political and cultural phenomenon than it is a technical one. Turning its framing over to scientists and other privileged experts, such as ethicists and social scientists, is to turn politics and culture over to them as well.

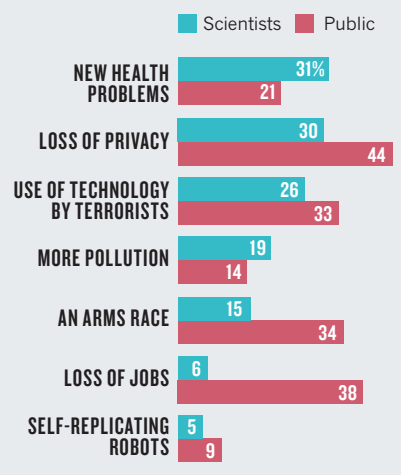
Scientists are not elected. They cannot represent the cultural values, politics and interests of citizens — not least because their values may differ significantly from those of people in other walks of life. A 2007 study<sup>6</sup> on the social implications of nanotechnology, for instance, showed that nanoscientists had little concern about such technologies eliminating jobs, whereas the public was greatly concerned (see 'A matter of perspective'). Each group was being rational. Nanoscientists have good reason to be optimistic about the opportunities created by technological frontiers; citizens can be justifiably worried that such frontiers will wreak havoc on labour markets.

Opening up questions of risk to democratic debate is on the whole good for science and innovation. The physicist Alvin Weinberg, a strong advocate for nuclear power, recognized this in the 1970s. Weinberg noted that the public debate of "questions like the probability of a reactor accident runs the risk of introducing exaggeration and distortion". Yet he also recognized that public pressure in the United States led to much greater attention to reactor safety than in the Soviet Union, where the public did not have a right "to participate in scientific and technological debate"<sup>7</sup>.

Different cultural and political approaches to choosing and managing risks invite different approaches to problem solving. Having

## A MATTER OF PERSPECTIVE

Surveys conducted in 2007 suggested that nanoscientists had different ideas from the general public about the main risks of nanotechnology.



rejected nuclear power, Germany is becoming a demonstration project for renewable-energy technologies, even as its neighbour France has shown how nuclear can provide an alternative low-carbon energy system. Opposition to GM was described as "a form of madness" by former European Commission science adviser Anne Glover, but it is part of a broader consumer movement that stokes demand for large-scale organic farming, integrated pest management, reduced use of antibiotics and reduced consumption of beef. Such preferences open up alternative innovation pathways that can add diversity and resilience to the global food system.

These ongoing debates show that the scientific community's efforts to wrest control over the specification of technological risk have not worked. Instead they have undermined the legitimacy of science.

As new areas of contentious technology emerge, the way out of this situation is to let democratic deliberation lead the way in determining which values and world views ought to be protected and which sacrificed.

## WORLDWIDE VIEWS

If an informed public discussion is needed, then let's have one. The capacity of people to learn about and deliberate wisely on the technical aspects of complex dilemmas has been documented by social scientists for decades<sup>8</sup>.

One model for how such discussions can be organized on an international scale has been developed by the World Wide Views (WWV) alliance, coordinated by the Danish Board of Technology. Since 2009, WWV has convened deliberations among diverse groups of about 100 citizens at numerous sites around the world — on global warming, biodiversity, and, earlier this month, climate and energy. Thousands of people

have participated from all corners of society. (In Washington DC, a WWV group discussing biodiversity included a homeless person, a roofer and a physicist.)

Each WWV deliberation is held across the world during a single day. Participants are provided with the same written and video background material (vetted by expert panels) on the issue being discussed. The day is divided into four or five thematic sessions; participants, in moderated groups of five to eight people, discuss a set of questions for each theme, then vote on relevant policy choices. It is too early to assess its actual impact on policy, but WWV demonstrates the viability of large-scale, representative deliberation on complex matters of global import.

Institutional models are also emerging that involve the public directly in choices about research that could influence the very nature of human existence — what might be termed the sciences of the existential. In the United States, for example, NASA last year commissioned the Expert and Citizen Assessment of Science and Technology (ECAST) network, to convene public deliberations on options for asteroid detection, mitigation and recovery. The results are informing agency decisions.

If the sciences of the existential are at hand, then let's make decisions about them collectively. WWV-type deliberations could address questions about what is acceptable and what isn't, about appropriate governance frameworks for research, and about the relative priority of different lines of study given ongoing and inevitable uncertainties and disagreements about risks and benefits.

This sort of discussion should continually feed into and set the boundary conditions for expert panels. A truly deliberative process that is geographically distributed and demographically inclusive can reveal the variations in how risks are selected and prioritized in different places and cultures. Values, governance regimes and research agendas can co-evolve in response to such knowledge. Democracy and science will both be better off. ■

**Daniel Sarewitz** is professor of science and society and co-director of the Consortium for Science, Policy and Outcome at Arizona State University, Tempe, Arizona, and is based in Washington DC, USA.  
e-mail: daniel.sarewitz@asu.edu

1. Liang, P. *et al.* *Protein Cell* **6**, 363–372 (2015).
2. Lanphier, E. *et al.* *Nature* **519**, 410–411 (2015).
3. Baltimore, D. *et al.* *Science* **348**, 36–38 (2015).
4. Hicks, D. J. *Stud. Hist. Phil. Biol. Biomed. Sci.* **50**, 1–12 (2015).
5. Marris, C. *et al.* *Public Perceptions of Agricultural Biotechnologies in Europe* (Lancaster Univ., 2001).
6. Scheufele, D. *et al.* *Nature Nanotechnol.* **2**, 732–734 (2007).
7. Weinberg, A. *Minerva* **10**, 209–222 (1972).
8. Peterson, J. C. (ed.) *Citizen Participation in Science Policy* (Univ. Massachusetts Press, 1984).

# Move beyond differences

Researchers and ethicists need to see past what can seem to be gendered debates when it comes to the governance of biotechnology, says **Charis Thompson**.

This autumn, researchers and other experts will come together to discuss the scientific, ethical and policy issues associated with gene-editing research in humans. Plans for the international meeting were announced by the US National Academy of Sciences and the National Academy of Medicine after a study was published in which researchers used a gene-editing tool known as CRISPR to modify the genomes of non-viable human embryos<sup>1</sup>.

Whether this meeting and others like it, planned in the United States, can help to forge a path for gene editing that takes into account all the relevant needs and concerns will depend on what efforts are made to integrate the diverse perspectives of people with different expertise and values. A first step to such integration is understanding how different perspectives arise.

One division in cares and concerns seems at times to fall along stereotypical gender lines. This was powerfully demonstrated during a meeting in Atlanta, Georgia, last month on biotechnology and ethics. About 200 global thought leaders gathered at BEINGS 2015 to “reach consensus on the direction of biotechnology for the twenty-first century”. On the day I attended, it was generally men who focused on containing biosecurity threats and on how to prevent regulation from impeding research. Women raised concerns about eugenics and class, race and gender inequalities in relation to biotechnology. Women were also the ones discussing the environment, the future of humanity, and the possible harms to the people who supply materials such as tissues and eggs, on which advances in biotechnology rely.

People at the conference quipped that this gendered divide reflected the following: men are pro-science and women are pro-ethics; men draw on rational criteria to support their arguments whereas women draw on emotional ones; men are interested in tangible, pragmatic issues whereas women are interested in values and deep ethical thought.

All of these ideas are reductive and sexist. They fail to recognize that people — whatever their gender, race or class — generally focus on pragmatic and measurable solutions to the problems they find the most pressing. People just differ on which issues they think are the most important.

That men and women at BEINGS tended to focus on different concerns reflected underlying differences of opinion about what needs to be taken into account when considering whether and how a particular area of science should proceed. These underlying differences are themselves gendered, because gender still affects which fields and subfields people enter, who does what kind of professional and domestic work, and where people’s political concerns lie<sup>2</sup>.

Take questions of framing, governance and jurisdiction. When framing the problems, people in one camp tend to see cures for disease as a morally unassailable good underlying the investment and effort currently being channelled into technological advances. The people in this camp generally believe that biotechnology (and science in general) should primarily be self-policing. They also tend to perceive non-experts and scholars with different areas of expertise as outsiders who slow down or stifle research.

For those in the other camp, the worth of scientific breakthroughs should be judged in a broader social context: addressing certain problems, such as entire populations being medically underserved, should be an integral part of the responsible development of biotechnology. People in this camp also tend to believe that the governance of biotechnology should be shaped according to how and where biotechnology is developed and deployed, for whom and at what cost — and that a greater diversity of experts, and many more non-experts, should be involved.

Obviously, there is not a clean two-gendered split, with men neatly falling into one camp and women into the other. Multiple genders, sexualities and family structures are increasingly being recognized worldwide, biologically and politically. Numerous scholars have shown that the

*“People generally focus on pragmatic and measurable solutions to the problems they find the most pressing.”*

effects of a person’s gender on their choice of work, politics and so on can be understood only if their class, race, sexuality, disability and citizenship are also considered (see, for example, ref. 3). Nonetheless, differences in approaches to bioethics come out in a gendered way because the factors that shape views of the appropriate scope and nature of science governance are themselves gendered.

Currently, there is a tendency for some scientists, bioethicists, lawyers, patient advocates and others to regard as ‘woolly’ people who are concerned with issues such as health disparities. Equally, those who take a broader view of biotechnology’s responsibility perceive some scientists as too caught up in the monetization and competition of their fields. These views often manifest as a perceived deficit in rationality in the other camp.

When scientific developments throw up difficult choices, scientists, social scientists and others need to stop positioning other people as being for or against science, or for or against ethics.

Mechanisms are needed to ensure that neither camp dominates the other. Following the autumn meeting, the National Academy should establish working groups to study and document at regular intervals the wide range of phenomena that will shape and be shaped by gene-editing research — from scientific breakthroughs to health disparities and disability justice.

If specialist pursuit of certain issues in such working groups is combined with efforts to bring the diverse strands of understanding together, for instance in the same publications, we might end up with better science and better ethics. ■

**Charis Thompson** is professor and chair in the Department of Gender and Women’s Studies, University of California, Berkeley, California, USA, and professor in the Department of Sociology, London School of Economics, UK. She is the author of *Making Parents and Good Science*.  
e-mail: [charis@berkeley.edu](mailto:charis@berkeley.edu)

1. Liang, P. et al. *Protein Cell* **6**, 363–372 (2015).
2. European Commission. *She Figures 2012: Gender in Research and Innovation — Statistics and Indicators* (Publications Office of the European Union, 2013).
3. Grzanka, P. R. (ed.) *Intersectionality: A Foundations and Frontiers Reader* (Westview, 2014).







The El Toro II glacier on the Chile–Argentina border, near the Pascua Lama gold mine.

## ENVIRONMENTAL POLICY

# Legislation left out in the cold

**Mark Carey** examines the cautionary tale of Argentina's struggle to pass the world's first glacier-protection law.

**T**he human-driven 'big melt' in the cryosphere is a major concern, with issues such as reduced glacier run-off riddling headlines. Glaciers are crucial, scientifically, socially and economically: they affect sea level, feed rivers, yield some of the world's most precise climate data in ice cores, and draw tourism to icy landscapes. But as Jorge Daniel Taillant explains in *Glaciers*, few people are actually doing anything about the melt — except in Argentina, where in 2010 authorities implemented what he refers to as the world's first glacier-protection law.

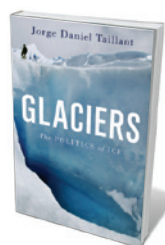
The rocky path towards creation of that law started in 2006, when Argentina's secretary of the environment (and Taillant's wife) Romina Picolotti learned that the Barrick Gold mining company of Toronto, Canada, had dynamited three glaciers at its Pascua Lama mine on the Chile–Argentina border to access a gold seam. What followed was a struggle to create and implement a national law to save glaciers from the damaging impacts of mining, hydrocarbon exploration and industrial development. Taillant

refers to the quest as cryoactivism, a term coined by French glaciologist Bernard Francou.

*Glaciers* is a

blow-by-blow account of progress towards the enactment and implementation of the law. It details both Picolotti's glacier-saving activism and Taillant's own, through the Center for Human Rights and Environment (CEDHA) — a non-governmental organization in Cordoba, Argentina, of which he is executive director. As a result, the book is largely an autobiographical account. But it offers discussion of glacier features, rock glaciers and the periglacial system, brief accounts of glacier shrinkage around the world and a chapter that includes an assortment of vignettes about human–glacier dynamics — from the creation of artificial glaciers in India to the 5,300-year-old mummy Ötzi, found in ice on the Austrian–Italian border.

Taillant notes that in 2008, Picolotti ushered the Minimum Standards Law for the Protection of Glaciers and the Periglacial Environment through Argentina's Congress with a unanimous vote in favour. The law sought to conserve glacial and periglacial



**Glaciers: The Politics of Ice**  
JORGE DANIEL  
TAILLANT  
Oxford Univ. Press USA:  
2015.

environments as hydrological resources, appeasing environmentalists and anti-mining activists who sought to save such resources for agriculture and potable water. But two weeks later, President Cristina Fernández de Kirchner vetoed the law.

Taillant speculates that "it is more than likely" that Barrick Gold president Peter Munk intervened. He bases his argument on three things: the law directly threatened billions of dollars of gold buried at Pascua Lama; Munk had "a one-on-one relationship with the Kirchners"; and Argentina's mining provinces (which might otherwise have been suspected of providing the opposition) remained supportive of the law. In 2010, several provinces enacted their own glacier-protection laws; Santa Cruz was "the first jurisdiction in the world" to adopt one. Later that year, environmental activists got support from Senator Daniel Filmus and Congressional Deputy Miguel Bonasso to push a new national glacier-protection law through the Congress: this one defined glaciers as "strategic freshwater reserves" for Argentina. Kirchner approved it.

Taillant shows how enforcement of this law has had a bumpy ride. A national glacier inventory, crucial for identifying what needs protection, was mandated but remains unfinished. In 2010, in an effort to stand up to what he sees as mining-company power,

PAV JORDAN/REUTERS/CORBIS

➔ **NATURE.COM**  
For more on science  
in culture, see:  
[nature.com/  
booksandarts](http://nature.com/booksandarts)

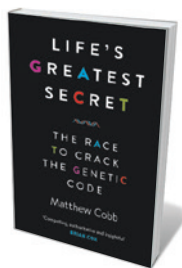
Taillant enrolled in a one-week glaciology course. He has since worked through CEDHA to identify glaciers using satellite imagery, and to conduct observations in mining regions. He strives to pinpoint glacier ‘hotspots’, where protection is lacking or mining claims or roads are a threat. Lower-level Argentinian courts have twice ruled to uphold the glacier law.

Although it is informative on Argentina’s environmental policy, Taillant’s book misses the mark on some of the science — especially when compared to glaciology texts or books such as Mariana Gosnell’s *Ice* (Univ. Chicago Press, 2007) and Mark Bowen’s *Thin Ice* (Henry Holt, 2005). When it discusses glaciers as water sources, the book does not precisely quantify the contributions of glacier runoff to downstream water supplies. It suggests that all glacial ice translates directly into water for human consumption, which exaggerates glacial input and runs contrary to studies that reveal that glaciers contribute up to 30% of downstream water in Peru and 27% in Bolivia, with sharp variations between dry and wet seasons. Nor does Taillant explain how groundwater, precipitation and land-use practices influence hydrology in glacierized watersheds. The book’s claim that perennial snow is the same thing as a glacier, and reference to “glaciers” that are only 3 metres thick (and therefore lack internal ice deformation, a defining characteristic) will raise eyebrows among glaciologists. When referring to diminishing glacier runoff in the Himalayas, the book cites treehugger.com — and at other points, Wikipedia.

Taillant argues that glacial contributions to hydrology are the basis for the glacier-protection law. But we do not know what proportion of Argentina’s downstream water they supply, during what season, or how that varies from glacier to sea. Without quantifying or fully explaining these contributions, the law and this book seem motivated by cultural values about endangered ice, rather than by scientific data. This tension is one that I have analysed before and that has been tackled by authors including Julie Cruikshank in *Do Glaciers Listen?* (Univ. British Columbia Press, 2005) and Eric Wilson in *The Spiritual History of Ice* (Palgrave Macmillan, 2009). Environmental laws, including those on glacier protection, evolve in the fraught national-to-global context of competing agendas, from the corporate to the governmental to that of environmental activists. As such, they frequently speak for, rather than with, local residents — who, as Taillant himself argues, have a human right to glaciers. ■

**Mark Carey** is an associate professor of history in the Clark Honors College at the University of Oregon in Eugene, and author of *In the Shadow of Melting Glaciers*. e-mail: [carey@uoregon.edu](mailto:carey@uoregon.edu)

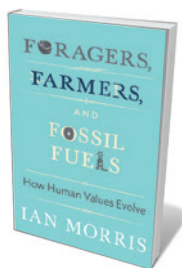
## Books in brief



### Life's Greatest Secret: The Race to Crack the Genetic Code

Matthew Cobb PROFILE (2015)

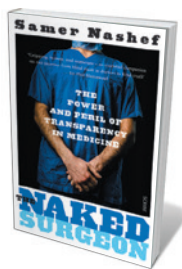
James Watson’s *The Double Helix* (Athenaeum, 1968) and Horace Judson’s *The Eighth Day of Creation* (Touchstone, 1979) are classics on the 1953 discovery of DNA’s structure, and the research that consolidated it. Zoologist Matthew Cobb richly recontextualizes the tale, tracing the interplay between biology, chemistry and physics that led to and amplified the breakthrough. This is a lucid explication of the science and the stories of key players, from X-ray crystallographer Rosalind Franklin and physicist Max Delbrück to Oswald Avery, who linked DNA to genes, and information theorist Claude Shannon.



### Foragers, Farmers, and Fossil Fuels: How Human Values Evolve

Ian Morris PRINCETON UNIVERSITY PRESS (2015)

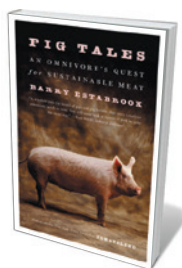
Energy capture — as a “brute material force” — shapes human values, argues archaeologist Ian Morris in this global-scale study of cultural variation. Looking at modes of capture from hunting and gathering to the agricultural and industrial revolutions, he avers that each determines population size, social organization and values. Now, writes Morris, our globalized society faces a “Malthusian collapse” — and a lurch towards new values — driven by nuclear weapons, climate change and digitization. With contributions from writer Margaret Atwood, sinologist Jonathan Spence and others.



### The Naked Surgeon

Samer Nashef SCRIBE (2015)

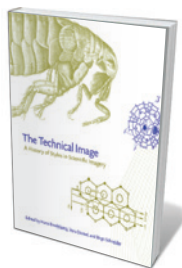
UK consultant cardiac surgeon Samer Nashef joins the swelling ranks of medics who have penned frank inside stories. Piquant detail abounds: operating on the heart, for instance, involves “twisting it, pressing on it, and sometimes turning it upside down”. But it is Nashef’s long study of risk that injects nuance. It began in 1977, when he discovered that arterial surgeons were responsible for the worst outcomes in a sample of abdominal aortic aneurysm operations. Such failures have, he shows, driven quality measurement in medicine, including his own heart-surgery risk model, EuroSCORE.



### Pig Tales: An Omnivore's Quest for Sustainable Meat

Barry Estabrook W. W. NORTON (2015)

Cheap pork should give us pause, writer Barry Estabrook argues in this cogent, level-headed investigation of the pig as raised and researched. The animal is one of the most intelligent ever domesticated, yet some 97 million in the United States are packed into reeking factory farms, with widespread ‘crating’ of pregnant sows. Yet as Estabrook reveals, humane, state-of-the-art husbandry is within our grasp and ultimately more profitable. He introduces an impressive group of animal scientists, farmers and pigs, from computer-savvy lab piglets to havoc-wreaking feral hogs.



### The Technical Image: A History of Styles in Scientific Imagery

Editors Horst Bredekamp, Vera Dünkel and Birgit Schneider UNIVERSITY OF CHICAGO PRESS (2015)

This multidisciplinary study trains an art historian’s eye on historical scientific imagery. Editors Horst Bredekamp, Vera Dünkel and Birgit Schneider draw on research from the Humboldt University of Berlin and a range of haunting images. They show that an iconic 1896 radiograph of a hand by X-ray discoverer Wilhelm Röntgen prompted both rhapsodies over a “photography of the invisible” and frustration among medics struggling to use such images for diagnosis. *Barbara Kiser*





In *Disaster Playground*, firefighters train for an asteroid impact.

# PLANETARY SCIENCE

## Space-rock alert

**Alexandra Witze** watches a pair of films on asteroids — according to many, a vast accident waiting to happen.

**S**tep aside, *Armageddon* and *Deep Impact*: two films are in competition to be this generation's seminal asteroid movie. Both offer crucial information about the asteroid threat to new audiences.

*Disaster Playground* and *51° North* are released widely on 30 June, the inaugural celebration of Asteroid Day — the anniversary of the 1908 meteorite explosion over Tunguska, Siberia. The event is meant to spread awareness of the threat of space rocks, repackaging the alarms that a small group of skywatching scientists have been issuing for decades. The novelty comes in how these films do it.

Both explore the ramifications of what might happen if an asteroid were discovered on a collision course with Earth. Both feature a narrator who interacts with real-life asteroid experts in a variety of offbeat settings. Both are a far cry from conventional documentaries, being more mash-ups of scientific information with fictionalized narratives.

*Disaster Playground*, directed by Nelly Ben Hayoun, is a visually arresting roller-coaster ride through California and the American Southwest, featuring interviews with high-impact scientists from NASA, the Sandia National Laboratories and elsewhere. It opens

### **Disaster Playground**

DIRECTOR: NELLY BEN HAYOUN  
Nelly Ben Hayoun Studio: 2014.

### **51° North**

DIRECTOR: GRIGORIJ RICHTERS  
Films United: 2015.

in Texas, with a cowboy and his horse startled by a huge blast. Ambulances and firefighters race through the debris left by the impact. Within minutes, meteor scientist Peter Jenniskens of NASA's Ames Research Center in Moffett Field, California, is presiding at an official-looking podium, sombrely reading a statement about the disaster.

In real life, Hayoun has the impressive title "designer of experiences at the SETI Institute", the organization in Mountain View, California, that explores life's place in the Universe. The film is certainly a designed experience. Hayoun ferries a bright-green toy dinosaur and a huge red telephone from interview to interview to represent the threat of mass extinction and the alerts needed to save the world from apocalypse. She uses offbeat camera angles, a pulsing soundtrack from rave band The Prodigy and title screens with giant capital letters to create a frenetic pace.

Hayoun coaxes scientists into the closest thing they may ever achieve to performance art. Ames astronomer David Morrison recreates his influential 1993 speech to Congress on the asteroid threat; his boss, Pete Worden, breaks out a Viking helmet and shield as a backdrop for an otherwise staid office interview. Hayoun's own appearances bind it all together as she hurries from shoot to shoot, pausing only to buy a pair of cowboy boots on her way to explore Disaster City, an emergency-response training centre in Texas, where search and rescue teams work their way through intentionally collapsed buildings. There is no time to wait when disaster strikes.

Grigoriy Richters' *51° North* is similarly kinetic but much more self-absorbed. It centres on a fictional YouTube star called Damon, who becomes obsessed with the asteroid threat. His reality-show audience is gradually alienated as he turns from taping his girlfriend and his dogs to making show after show about space-rock disasters. That all changes when — well, no spoilers. But given the title and the fact that it is about asteroids, you can guess what happens to London.

Richters has created a true millennial approach to the asteroid threat. Damon is constantly on camera: taping, uploading, tweeting, instagramming, sharing every moment. Brian May, Queen guitarist and astrophysicist, provides the soundtrack.

There is some science. Damon interacts with many of Britain's leading asteroid experts, and even tours the Spaceguard Centre in Knighton, the United Kingdom's leading asteroid-information centre. He works in smooth explanations of the various strategies for diverting an asteroid on a collision course with Earth. Footage from the many Russian car cameras that recorded the Chelyabinsk meteorite strike in February 2013 fits the video-focused approach. Unlike in *Disaster Playground*, there is no wider discussion of what the end of the world means for anyone other than the self-obsessed Damon.

Today, the cutting edge of asteroid science lies in pushing the limits of detection, searching for smaller and smaller risky space rocks. Scientists have pretty much nailed the chances of spotting any world-ending asteroid, a rock big enough to wipe out an entire region. What is left are the ones that can slip through the current net of telescopes and hit with little to no warning.

As New Year celebrations welcomed in 2014, an asteroid a few metres wide entered the atmosphere and disintegrated over the Atlantic Ocean — all while many of the astronomers who would normally track it were on a holiday break. We do live in a disaster playground. ■

**Alexandra Witze** is a correspondent for Nature based in Boulder, Colorado.

# Correspondence

## Russian science loses to politics

Philanthropist Dmitry Zimin is closing down his successful Dynasty Foundation — modern Russia's first private science-funding organization — after the Ministry of Justice fined it for being a “foreign agent” (see *Nature* **521**, 273; 2015). Like many other Russian scientists, we believe that these events will have dire immediate and long-term consequences for Russia's science.

Ironically, the government has been trying to revitalize Russian science in the past few years. Along with increased research funding to universities and a drastic overhaul of science management, it set up a mega-grant programme to attract back Russian scientists working abroad. It also created innovation centres and transferred fund management from the previously independent Russian Academy of Science to a government organization (see [go.nature.com/m75bj4](http://go.nature.com/m75bj4)).

The controversial law that claimed the Dynasty Foundation is intended to curtail perceived foreign influence in Russian politics. This case suggests that the future of Russian science depends on political forces to a greater extent than the government seems prepared to acknowledge.

**Fyodor A. Kondrashov**  
*University of Pompeu Fabra, Barcelona, Spain.*

**Alexey S. Kondrashov**  
*University of Michigan, Ann Arbor, USA.*

**Mikhail S. Gelfand** *Russian Academy of Sciences, Moscow, Russia.*  
[fyodor.kondrashov@crg.eu](mailto:fyodor.kondrashov@crg.eu)

## India: multi-author papers skew ranking

Your ranking of India's top ten institutions, scored by the number of research papers in the Scopus citation database over the past five years, is distorted by the

exceptionally large number of citations attracted by papers with hundreds of authors (see *Nature* **521**, 142–143; 2015).

This distortion is often corrected by limiting citations to individual authors — for example, by using an individual *h*-index for reference. Another corrective measure would be to exclude papers with more than 100 authors from citation metrics. For example, the Large Hadron Collider's ATLAS collaboration includes thousands of authors, so exclusion would have roughly the same effect as using the individual *h*-index (J. E. Hirsch *Proc. Natl Acad. Sci. USA* **102**, 16569–16572; 2005).

Exclusion would introduce major changes in some of your listed entries. For example, I calculate that about 20% of the publications attributed to the highly rated Panjab University have long author lists and contribute almost two-thirds of the citations. Excluding these papers reduces Panjab University's citation impact ratio from 1.4 to 0.7, causing it to drop out of your top ten; the fall is comparable for your second-placed Tata Institute of Fundamental Research in Mumbai. The fall for the Indian Institute of Technology at Bombay and Guwahati is less marked at 15% and 12%, respectively, because these each generate proportionally fewer papers with long author lists.

**P. Sriram** *Indian Institute of Technology, Chennai, India.*  
[sriram@iitm.ac.in](mailto:sriram@iitm.ac.in)

## India: assess social impact of technology

Researchers' priorities for improving science in India should include a commitment to assess the social impacts of new technologies in the Indian context (see *Nature* **521**, 151–155; 2015).

Big dams and atomic-energy programmes offered solutions to many of India's problems after independence in 1947; the green

revolution and biotechnology followed. Stem-cell therapy, nanotechnology, synthetic biology and pharmacogenomics are all now taking off.

India sometimes seems prepared to overlook the potential societal consequences of such technologies in the name of development and progress. Its Land Acquisition Bill 2015, for example, seeks to exempt some important projects on defence, infrastructure and industrial regions from social assessment.

The impact of new technologies on India's sizeable poor and vulnerable population should be analysed before such innovations are introduced (see D. Greenbaum *Nature Biotechnol.* **33**, 425–426; 2015). Analysis would need to include investigation of their affordability and equitability (S. S. Tiwari and S. Raman *New Genet. Soc.* **33**, 413–433; 2014).

**Shashank S. Tiwari** *ITT Labs, Birmingham, Alabama, USA.*  
**Ekta Tiwary** *University of Alabama at Birmingham, USA.*  
[shashank17t@gmail.com](mailto:shashank17t@gmail.com)

## Speed translation of misconduct reports

Two reports of scientific misconduct were publicly released around 19 May this year, relating to work by graduate student Michael LaCour at the University of California, Los Angeles, and by surgeon Paolo Macchiarini at Stockholm's Karolinska Institute (see *Nature* **521**, 406–407; 2015). Surprisingly, the unofficial report on LaCour's work spread like wildfire across the media, whereas the official report on Macchiarini's received relatively little attention.

This discrepancy could simply reflect popular interest in the different research areas investigated. But it is notable that the report on LaCour's work was in English ([go.nature.com/5dyi6w](http://go.nature.com/5dyi6w)), whereas that on Macchiarini's was first released in Swedish ([go.nature.com/bzwaxt](http://go.nature.com/bzwaxt)).

In my view, misconduct reports would be more useful if they were

simultaneously translated into English, the common language of scientific discourse. This would allow the wider scientific community to immediately assess the validity of the claims and to avoid new research based on falsified findings.

Of the European non-English-speaking research integrity offices, only the Danish Committees on Scientific Dishonesty report in English. This hastens the all-important correction of the scientific record by enhancing accessibility and transparency.

**Chris H. J. Hartgerink** *Tilburg University, the Netherlands.*  
[c.h.j.hartgerink@uvt.nl](mailto:c.h.j.hartgerink@uvt.nl)

## US sanctions alarm physicians from Iran

Iranian physicians in North America faced an unsettling constraint this month. Between 29 May and 5 June, the US Educational Commission for Foreign Medical Graduates (ECFMG) stopped processing requests to verify credentials issued by Iranian institutions, pending clarification of restrictions on interactions between US and Iranian medical and educational organizations (see [go.nature.com/jrcxzz](http://go.nature.com/jrcxzz)). The Medical Council of Canada, which also uses the ECFMG, made a similar announcement.

Iranian citizens are among the top ten groups of international physicians who acquired US certification in 2009 (see [go.nature.com/ecmnkb](http://go.nature.com/ecmnkb)) and were the largest group among immigrants entering post-medical training in Canada in the same year (see [go.nature.com/95ayyf](http://go.nature.com/95ayyf)).

As Richard Nephew — a former senior US official and leading sanctions architect — has warned, there is a risk that the US foreign-policy tool of using sanctions to isolate Iran could backfire (see [go.nature.com/ocrini](http://go.nature.com/ocrini)).

**Mehdi Aloosh** *Montreal, Canada.*  
[md\\_aloosh@hotmail.com](mailto:md_aloosh@hotmail.com)



# John Forbes Nash

## (1928–2015)

Master of games and equations.

John Forbes Nash, an exalted mathematician whose life took dramatic turns between genius, mental illness and celebrity status, made major contributions to game theory, geometry and the field of partial differential equations.

Nash, who died on 23 May, was born in Bluefield, West Virginia, in 1928. His father was an electrical engineer and his mother a schoolteacher. In 1945, after excelling in mathematics at high school, he attended the Carnegie Institute of Technology (now Carnegie Mellon University) in Pittsburgh, Pennsylvania. At first he studied chemical engineering, but soon after enrolling he switched to chemistry and then to maths.

In Nash's final year, one of his professors wrote a recommendation letter for the 19-year-old supporting his application to graduate school. It simply stated: "He is a mathematical genius." In 1948, Nash was accepted by Harvard University in Cambridge, Massachusetts, and by Princeton University in New Jersey. He chose Princeton.

As a PhD student, Nash proved the existence of the equilibrium that now carries his name. His 1950 paper 'Equilibrium points in  $n$ -person games', contains about 330 words, two references and not one equation (J. F. Nash Jr *Proc. Natl Acad. Sci. USA* **36**, 48–49; 1950). One of the citations is the 1944 book *Theory of Games and Economic Behavior* — in which mathematician John von Neumann and economist Oskar Morgenstern introduce game theory, a mathematical approach for studying strategic and economic decisions.

The Nash equilibrium is a position in a game from which none of the players can change their strategy to improve their pay-off. Imagine a game with two players (yourself and another person) and two strategies, A and B. If you both choose A, your pay-off is 2. If you choose A and your opponent chooses B, you score 0. If you choose B and the other player chooses A, your pay-off is 3. If you both choose B, you score 1. The same applies to your opponent.

In this example, the Nash equilibrium occurs when both players choose B. If both players choose B, their pay-off is 1; if either player switches to A, their pay-off falls to 0. In other words, neither player can independently switch their strategy and improve their pay-off. Observe that if both players select A, there is no Nash equilibrium because you could improve your pay-off by switching to B.

Calculating the Nash equilibrium can be



a formidable task in a complex game. There is also the uncertainty over whether the person you are playing against is sufficiently rational to play the equilibrium strategy. If both players are rational and their rationality is common knowledge, they would play it. But experiments often reveal that people are not rational. Regardless of whether people actually play the Nash equilibrium in social or economic interactions, working out what it is (or what the Nash equilibria are) is the first step to analysing any game.

Although dismissed at the time by von Neumann as a triviality, the Nash equilibrium has been used to analyse all sorts of competitive situations. As well as being key to decision-making in economics and politics, the idea is important in biology. Here, the nearly equivalent concept, formulated by evolutionary biologist John Maynard Smith in the 1970s is called an evolutionarily stable strategy (ESS). If all members of a population adopt an ESS, then natural selection prevents a rare mutant from spreading.

On completing his PhD, Nash joined the Massachusetts Institute of Technology (MIT) in Cambridge in 1951. He worked — first as an instructor and later as a professor — in the mathematics faculty until he resigned in 1959. It was while he was at MIT that Nash met and married Alicia Lopez-Harrison de Lardé, a physics student there.

Among mathematicians, Nash is best known for his work in real algebraic geometry

and nonlinear partial differential equations. He was not afraid to tackle the hardest problems in the field, and he succeeded. In 1957, he — in parallel with Italian mathematician Ennio de Giorgi — solved Hilbert's nineteenth problem involving partial differential equations.

It was during a talk in 1959 on what is seen to be one of the hardest problems in maths — the Riemann hypothesis — that the audience realized that there was something wrong with Nash. His talk was incomprehensible.

He was diagnosed with paranoid schizophrenia that year. Over the next two decades, Nash was in and out of hospitals. He underwent therapy, and for a while left the United States and sought asylum in Switzerland in an attempt to escape his imagined tormentors. For many years he wandered around the Princeton campus. Throughout this period, Alicia, who divorced Nash in 1963, oversaw much of his care.

In the late 1980s, Nash reappeared in academic circles, and in 1994 he was awarded the Nobel Memorial Prize in Economic Sciences for his work on game theory. The Nobel and the 2001 film *A Beautiful Mind*, based on journalist Sylvia Nasar's book of the same name, which recounted Nash's struggles, propelled him into the limelight.

In May this year, Nash received the Abel prize from the Norwegian Academy of Science and Letters for his work on partial differential equations. On the way back from the celebration in Norway, John and Alicia (who had remarried in 2001) were killed in a car accident in a taxi on the New Jersey turnpike. John was 86.

I met John in 1998 at the Institute for Advanced Study in Princeton. Over the years, I gave several talks there that he attended. One summer's day, when the usual sitting arrangements for lunch were disrupted by the closure of the main kitchen, I noticed John, the physicist Edward Witten and Andrew Wiles, the British mathematician who proved Fermat's last theorem, sitting down together at a small table. I wondered which of them would start the conversation. None of them did. I seem to remember that they ate their meal in silence. ■

**Martin A. Nowak** is professor of mathematics and biology, and director of the Program for Evolutionary Dynamics at Harvard University in Cambridge, Massachusetts, USA.  
e-mail: martin\_nowak@harvard.edu

➔ **NATURE.COM**  
For more on John Nash's life, see:  
[go.nature.com/bi7gxc](http://go.nature.com/bi7gxc)

## ASTROPHYSICS

## Dust-poor galaxies at early times

Observations of galaxies that formed early in the Universe's history reveal much lower dust levels than are found in sources from a slightly later era. It seems that galaxies underwent rapid change during a relatively short period. [SEE LETTER P.455](#)

VERONIQUE BUAT

The study of the most distant galaxies, observed as they were about 1 billion years after the Big Bang, is crucial for our understanding of the star-forming activity and the physical processes at work in these young systems. Writing in this issue (page 455), Capak *et al.*<sup>1</sup> present a study of nine such galaxies using a linked-up telescope array. They find that the dust and gas properties in these systems hint at an interstellar medium (ISM) that is much less evolved than in galaxies about 2 billion years older. This suggests that there was a rapid change in the overall properties of galaxies during the early life of the Universe.

The expansion of the Universe shifts the ultraviolet (UV) light emitted by newly formed stars in remote systems to longer (visible and near-infrared) wavelengths that, unlike UV light, can be observed by ground-based telescopes. The most distant objects known today are detected as a result of a break in the continuum of their redshifted spectra at wavelengths of around 0.1 micrometres; this is due to the absorption of UV photons by neutral hydrogen in the intergalactic medium. The absorption occurs for photons with energies corresponding to wavelengths shorter than the Lyman- $\alpha$  line of hydrogen (1,216 nm), and galaxies whose distances have been estimated by this method are known as Lyman break galaxies (LBGs). The most comprehensive surveys undertaken so far have led to detections of very young LBGs that formed approximately 0.5 billion years after the Big Bang<sup>2</sup>.

The presence of interstellar dust complicates the study of galaxies, and affects measurements of fundamental, observationally derived properties such as the star-formation rate. This is because dust is efficient at absorbing the energetic UV photons (a proxy for the star-formation rate) that are emitted by young stars and at re-emitting their energy in the infrared domain, at wavelengths longer than 5  $\mu\text{m}$ . This is a complex process that depends not only on the amount of dust present, but also on its distribution relative to the stars and on its composition<sup>3</sup>. Overall, dust substantially reduces the intensity of stellar light reaching the telescopes<sup>4</sup>.



ESO/C. MALIN

**Figure 1 | The Atacama Large Millimetre Array (ALMA).** Capak *et al.*<sup>1</sup> used 20 of ALMA's antennas to study the interstellar medium (ISM) of 9 galaxies that were present when the Universe was only about 1 billion years old. The authors found that their sources contain a smaller amount of dust than expected. Some of the galaxies in the sample may have an ISM similar to that of the Small Magellanic Cloud (a satellite galaxy of the Milky Way), which is visible here (right of centre) as the smaller of the two Magellanic Clouds above the antennas.

A straightforward method to account for the UV light produced in galaxies involves observing the radiant energy that is absorbed by dust, is re-emitted and is then redshifted in the far-infrared and submillimetre domains. However, the low sensitivity of detectors, combined with the poor spatial resolution achieved by single-dish telescopes, make surveys of high-redshift galaxies at these wavelengths less efficient than those at optical or near-infrared wavelengths. Even the Herschel Space Observatory, which detected<sup>5</sup> the infrared emission from dust in galaxies at redshifts of up to 2–3 (corresponding to a time roughly 2 billion to 3 billion years after the Big Bang), was able to detect only hyper-luminous sources at much larger distances<sup>6</sup>.

Given that directly measuring the long-wavelength emission from dust is so challenging, astronomers resort to empirical relations

to derive dust's infrared luminosity. One such relation links this luminosity to the stellar UV luminosity<sup>7</sup> for a representative sample of nearby, actively star-forming galaxies, for which both luminosities have been accurately measured. Unfortunately, this recipe is not universally applicable because it depends on the properties of the ISM (such as the composition of dust and its distribution relative to the stars), as well as on the stellar populations in the galaxies<sup>8</sup>. Despite these caveats, however, it is extensively used to estimate the level of obscuration of the stellar UV light by ISM dust for galaxies across a wide redshift range. It will therefore be important to check the validity of this relationship — especially for young, high-redshift systems. A critical evaluation could also yield clues to the properties of the ISM at those early times.

Capak *et al.* used the Atacama Large Millimetre Array (ALMA), which was designed to



overcome both the resolution and sensitivity problems (Fig. 1). Being an interferometer (a series of telescopes linked up to combine astronomical observations), ALMA has a small field of view that is suitable for observing well-centred sources, and it can detect the weak submillimetre emission originating from dust in ordinary galaxies at high redshifts. The authors used 20 of ALMA's antennas in unison to observe the dust and gas emissions of 9 typical LBGs located at redshifts 5–6; these correspond to a time when the Universe was about 1 billion years old.

Capak and colleagues selected their sample from the Cosmic Evolution Survey field, a two-square-degree area that has been extensively observed by most of the major telescopes, from the ground and from space. ALMA detected the thermal dust emission in four galaxies, and an ISM spectral line emitted from gaseous carbon at a wavelength of 158  $\mu\text{m}$  in all nine of them; the carbon feature is the dominant ISM emission line of galaxies in the far-infrared domain. Such a high detection rate is outstanding, because previous attempts failed to simultaneously detect the carbon feature and thermal dust emission<sup>9</sup>.

The authors' study argues for a very low dust content and stellar-light obscuration in these systems. The four galaxies whose thermal dust emission was detected may harbour an ISM similar to that of the Small Magellanic Cloud (a satellite galaxy of the Milky Way), which is characterized by a low abundance of elements heavier than helium. The upper limits put on the dust emission of the remaining five sources call for an even more extreme situation with a much lower infrared emission. That seems to be at odds with the observed UV luminosity of these systems. The enhanced carbon emission-line intensities also suggest low dust levels relative to the gas present in these early galaxies, although other explanations cannot be excluded.

An immediate consequence of these findings is that the classical calculations used to derive obscuration due to dust from the observed UV continuum luminosity are unlikely to be valid for LBGs in the early Universe. As a result, the star-formation rate considered to be appropriate for these galaxy types is likely to have been overestimated by factors of between two and four in previous studies. Last but not least, this

pioneering work paves the way for future observational campaigns. Although observing the low levels of dust emission from large samples of high-redshift galaxies may prove challenging even for ALMA, Capak and co-workers' finding of enhanced carbon emission lines should become a useful tool in the study of star-forming galaxies at those early epochs. ■

**Veronique Buat** is at the *Laboratoire d'Astrophysique de Marseille, Marseille 13013, France.*

*e-mail: veronique.buat@lam.fr*

1. Capak, P. L. *et al. Nature* **522**, 455–458 (2015).
2. Bouwens, R. J. *et al. Astrophys. J.* **803**, 34 (2015).
3. Witt, A. N. & Gordon, K. D. *Astrophys. J.* **528**, 799 (2000).
4. Burgarella, D. *et al. Astron. Astrophys.* **554**, A70 (2013).
5. Gruppioni, C. *et al. Mon. Not. R. Astron. Soc.* **436**, 2875–2876 (2013).
6. Riechers, D. A. *et al. Nature* **496**, 329–333 (2013).
7. Meurer, G. R., Heckman, T. M. & Calzetti, D. *Astrophys. J.* **521**, 64 (1999).
8. Boquien, M. *et al. Astron. Astrophys.* **539**, A145 (2012).
9. Maiolino, R. *et al. Mon. Not. R. Astron. Soc.* (in the press); Preprint at <http://arxiv.org/abs/1502.06634v2>

## NEURODEGENERATION

# Evolved protection against human prions

**A genetic variant of PrP, the protein that forms prions, confers protection against the human prion disease kuru by inhibiting the conversion of functional isoforms to the abnormal, disease-causing conformation. SEE LETTER P.478**

GLENN TELLING

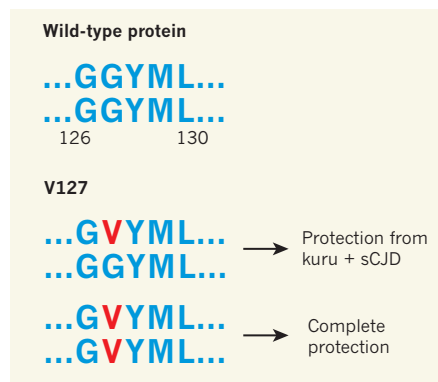
The group of neurodegenerative disorders known as transmissible spongiform encephalopathies are incurable and lethal. They are caused by prions, a type of infectious agent that has challenged our understanding of the concepts of infection and inheritance. In prion diseases, an abnormal isoform of prion protein (PrP), dubbed PrP<sup>Sc</sup>, induces conformational change in the functional isoform PrP<sup>C</sup>. The functional form adopts a disease-causing structure, and these abnormal proteins propagate throughout the brain, damaging nerves. But the mechanistic underpinnings of this infectious propagation remain enigmatic. On page 478 of this issue, Asante *et al.*<sup>1</sup> describe a form of PrP<sup>C</sup> that seems to be resistant to such conformational change, providing some insight into the long-standing issue of whether it is possible to prevent prion diseases.

Kuru, a human prion disease, became a

leading cause of death among the Fore population of the highlands of Papua New Guinea in the mid-twentieth century<sup>2</sup>, killing more than 3,000 people in a population of 30,000. The aetiology of the disease was initially puzzling, but in 1966, a study<sup>3</sup> showed that brain extracts from people who had kuru caused progressive neurodegeneration when introduced into chimpanzees, following a prolonged incubation period. This supported the idea that kuru is an infectious disorder. In fact, kuru is a devastating consequence of ritualistic cannibalism, in which the brains and other body parts of tribal members were consumed as an act of mourning.

Fast forward to the end of the twentieth century, and interest in the kuru epidemic was renewed in response to the emergence of variant Creutzfeldt–Jakob disease (vCJD). This human prion disease results from exposure to prions that cause bovine spongiform encephalopathy, carried by cows with ‘mad cow’ disease. Then, in 2009, the authors of the current study discovered a genetic variant (a polymorphism) in *PRNP*, the gene that encodes PrP (ref. 4). The variant, which changes the amino acid at residue 127 from glycine (G) to valine (V), was present in one of the two copies of *PRNP* (mammals have two copies of each chromosome) in unaffected individuals in the kuru-exposed population. The authors proposed that the variant, called V127, conferred resistance to prion disease, and that it had been selected for in response to the kuru epidemic.

Asante *et al.* set out to compare V127 with another polymorphism in close proximity



**Figure 1 | Valine variants.** The wild-type amino-acid sequence of human prion-protein residues 126 to 130 is glycine (G), G, tyrosine (Y), methionine (M) and leucine (L). Replacement of G for V at residue 127 is known to confer protection against the prion disease kuru. Asante *et al.*<sup>1</sup> demonstrate that, in mice, one copy of V127 also confers resistance to sporadic Creutzfeldt–Jakob disease (sCJD), and two copies protect against all human prion diseases.

— amino-acid residue 129 of PrP, which can be either methionine (M) or V. It is thought that one copy of each 129 variant confers some resistance to human prion diseases by affecting the binding between identical prion proteins that is required for conformational change. The researchers injected human prions into mice engineered to express human forms of PrP harbouring different combinations of the two genetic variants.

Mice with one copy of V127 were completely protected against kuru prions, and died of old age without developing neurological disease. By contrast, all mice without V127 developed disease after around 200 days, regardless of the variant at residue 129, confirming that this model system accurately recapitulates the resistance to kuru conferred by V127. Previous work<sup>5</sup> has indicated that the transmission properties of kuru are similar to those of the sporadic form of CJD (sCJD), and so it probably came as no surprise to the authors to find that mice with one copy of V127 were also protected from sCJD prions. Interestingly, however, not all of these mice were protected against vCJD prions.

Remarkably, mice carrying two copies of V127 were completely protected against all forms of human prion disease (Fig. 1), indicating that the polymorphism at residue 127 acts in a different manner from that at 129. These mice failed to manifest any signs of prion disease, a noteworthy response that was mimicked only in mice that did not express any prion protein. Finally, Asante *et al.* investigated how the protective effect of one copy of V127 varied when wild-type PrP (in which the genetic sequence includes the G127 and M129 variants) was expressed at different levels. These experiments demonstrated that V127 acts as a 'dominant negative' inhibitor of prion conversion — not only is it itself resistant to conformational conversion, but it also inhibits conversion of wild-type proteins.

It will be important to learn what structural consequences, if any, occur when V replaces G at residue 127. There are currently no treatments or cures for transmissible spongiform encephalopathies, and structural information may hint at ways to inhibit PrP<sup>Sc</sup> propagation and maintain normal PrP<sup>C</sup> function. Approaches that capitalize on the profound dominant-negative properties of V127 are obvious targets for further investigation.

V127 was naturally selected in response to epidemic human prion disease. In a similar manner to DNA- or RNA-based infectious agents, prions are subject to Darwinian forces, such that the structure of the disease-causing protein can be altered through selective pressure<sup>6</sup>. This begs the question of whether new prion strains might emerge to counter the protective effects of the V127 variant. It will be interesting to determine whether V127 also prevents conformational changes in PrP

in experimental models of prion amplification, which have been shown to enable interspecies prion transmission<sup>7</sup>, and if so, what the infectious properties of the resulting prions might be. Finally, it will be valuable to investigate the V127 polymorphism in the context of PrP from other species that are naturally susceptible to prion diseases, using similar approaches to those in the current study. This should help to ascertain whether the effects of this fascinating variant are universally protective against all combinations of species and prion strains. ■

**Glenn Telling** is at the Prion Research Center, Colorado State University, Fort Collins, Colorado 80525, USA.

e-mail: glenn.telling@colostate.edu

1. Asante, E. A. *et al.* *Nature* **522**, 478–481 (2015).
2. Hadlow, W. J. *Lancet* **274**, 289–290 (1959).
3. Gajdusek, D. C. *Nature* **209**, 794–796 (1966).
4. Mead, S. *et al.* *N. Engl. J. Med.* **361**, 2056–2065 (2009).
5. Wadsworth, J. D. *et al.* *Proc. Natl Acad. Sci. USA* **105**, 3885–3890 (2008).
6. Li, J. *et al.* *Science* **327**, 869–872 (2010).
7. Vidal, E. *et al.* *J. Neurosci.* **33**, 7778–7786 (2013).

This article was published online on 10 June 2015.

## PLANT SCIENCE

# Precision positioning with peptides

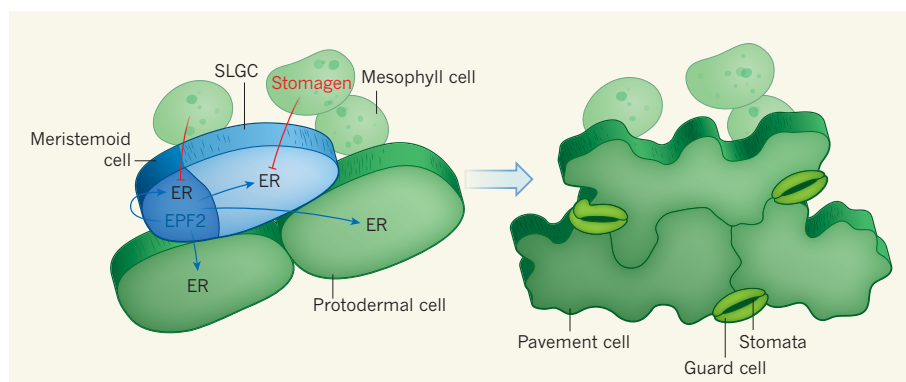
Two related peptides compete for binding to the same receptor to regulate the spacing of cells on the lower surfaces of leaves. This discovery highlights the complexity of cell signalling in plants. [SEE ARTICLE P.439](#)

## SACCO DE VRIES

Being fixed in position, plants must be able to respond to a multitude of environmental and developmental inputs over appropriate time frames. To do this, they use organic molecules such as auxin or ethylene, as well as signalling molecules analogous to those used in animals, including peptides (short strings of amino acids) and hormones. But plant peptide signalling pathways, particularly interactions between peptides and their receptor proteins, are poorly understood. In

this issue, Lee *et al.*<sup>1</sup> (page 439) define a plant signalling pathway in which two peptides compete for binding to one receptor, exerting opposing effects on the development of tiny openings called stomata on the lower surfaces of leaves.

Stomata regulate the passage of oxygen, carbon dioxide and water vapour in and out of the leaf. These pores are surrounded by two guard cells, which mediate stomatal opening and closing in response to the time of day, to changes in the environment or to physical changes in the plant itself. Guard cells originate



**Figure 1 | Orchestrating stomatal development.** During leaf development, most protodermal cells develop into pavement cells, which line the underside of leaves. But a few can undergo differentiation and asymmetric division to form meristemoid cells and stomatal lineage ground cells (SLGCs), which together give rise to the guard cells that enclose stomata — pores on the surface of the leaf. Meristemoids, SLGCs and protodermal cells all express the receptor protein ERECTA (ER). ER activation by the peptide EPF2, which is secreted from cells that have already become meristemoids, suppresses the meristemoid–guard-cell lineage in adjacent cells. By contrast, the peptide Stomagen, which is produced in the overlying mesophyll cell layer, inhibits activation of ER and so promotes guard-cell differentiation. Lee *et al.*<sup>1</sup> report that EPF2 and Stomagen compete for binding to ER, to regulate the even spacing of guard cells between pavement cells.



from precursors called protodermal cells. Protodermal cells can either proliferate to give rise to pavement cells, which make up the protective epidermis covering the surface of the leaf, or differentiate into a meristemoid mother cell (MMC), a progenitor of the guard cell. MMCs divide asymmetrically, one daughter cell becoming a meristemoid, the other becoming a stomatal-lineage ground cell, a progenitor that can either differentiate into pavement cells or, through a 'spacing division', give rise to a second, satellite meristemoid cell. Meristemoids then undergo asymmetric divisions, and finally differentiate into pairs of guard cells that are evenly distributed between pavement cells (Fig. 1). But what governs these cell-fate decisions?

Mutations in genes encoding receptor proteins of the ERECTA (ER) family cause a loss of control over stomatal spacing, leading to the development of leaves with unevenly spaced stomata. The severity of this trait is further influenced by mutations in an ER co-receptor, TOO MANY MOUTHS (TMM), demonstrating that ER-family signalling is a key regulator of stomatal development. ER-family receptors are activated by EPIDERMAL PATTERNING FACTOR 1 (EPF1) and EPF2, peptides that exert different, but overlapping, effects on stomatal development. By signalling through the receptor ER-LIKE 1 (ERL1), EPF1 orients stomatal spacing and prevents differentiation into guard cells, whereas EPF2-ER signalling restricts the formation of this lineage at an earlier stage. Both receptors exert their effects by activating an intracellular signalling cascade called the MAPK pathway (reviewed in ref. 2).

In contrast to EPF1 and EPF2, the EPF-like peptide Stomagen promotes stomatal development by inducing guard-cell differentiation<sup>3</sup>. Lee *et al.* set out to investigate how these related peptides could exert such dramatically opposing effects. The authors found that the altered stomatal differentiation that occurs in plants harbouring mutant forms of the genes encoding ER, ERL1 or TMM was not affected by the addition of Stomagen. This suggests that Stomagen, like EPF1 and EPF2, acts through ER-family receptors.

Lee and colleagues then provided extensive genetic evidence to support this idea. By inhibiting or overexpressing the gene encoding Stomagen, the authors showed that misexpression of the peptide interferes with all signalling pathways mediated by the ER family. Next, using sophisticated genetic and *in vitro* biochemical experiments, they confirmed that Stomagen actively competes with EPF2 for ER, the two peptides binding to the receptor with similar affinities. Finally, the researchers showed that Stomagen could not activate MAPK signalling. They propose that binding of Stomagen to ER-family receptors actually prevents MAPK signal transduction.

These results provide evidence that a finely balanced system of closely related activating and inhibitory peptides locally modulates the signalling pathways required for proper stomatal spacing in the plant epidermis (Fig. 1). One uncertainty in this model is the similarity in the dissociation constants (a measure of the binding strength between two proteins) of EPF2-ER and Stomagen-ER. This implies that massive changes in the local concentration of one ligand, or a large adjustment in the ligand's local receptor sensitivity, would be needed to induce a change in stomatal patterning. Which mechanism is used in physiological conditions is an issue that clearly needs to be addressed.

A competitive inhibition mechanism provides precise control of cellular spacing, analogous to the 'lateral inhibition' mechanism that regulates the spacing of animal cells. It is now imperative to determine whether similar mechanisms operate over time to generate repetitive differentiation events periodically during plant development. For example, systems such as the CLE40-CLV1-ACR4 signalling pathway, which regulates stem-cell maintenance in the growing tips and roots of plants<sup>4</sup>, might use a similar combination of activators and inhibitors. Future studies should also analyse whether local changes in the levels of particular peptide

combinations could weave a 3D pattern of locally activating and inhibitory signalling conditions.

The results of the current study can be put in the larger context of plant signalling pathways — which includes different combinations of ligands, receptors and co-receptors — acting in conjunction with local changes in signalling pathways mediated by organic molecules. Furthermore, because there is evidence for a direct connection between plant steroid signalling and the stomatal pathway<sup>5</sup>, yet another level of signalling seems plausible. A picture emerges in which almost every cell in a plant tissue can be exposed to a precisely defined set of signals. This new concept in plant spatial control should now be tested in other settings. ■

**Sacco de Vries** is in the Department of Biochemistry, Wageningen University, Wageningen 6703 HA, the Netherlands. e-mail: sacco.devries@wur.nl

1. Lee, J. S. *et al.* *Nature* **522**, 439–443 (2015).
2. Pillitteri, L. J. & Torii, K. U. *Annu. Rev. Plant Biol.* **63**, 591–614 (2012).
3. Sugano, S. S. *et al.* *Nature* **463**, 241–244 (2010).
4. Stahl, Y. *et al.* *Curr. Biol.* **23**, 362–371 (2013).
5. Gudesblat, G. E. *et al.* *Nature Cell Biol.* **14**, 548–554 (2012).

This article was published online on 17 June 2015.

## CLIMATE SCIENCE

# The dynamics of temperature extremes

**Changes in the occurrence of atmospheric circulation patterns are not well understood. A study finds that these have been a big factor in observed changes in regional temperature extremes during recent decades. SEE LETTER P.465**

**THEODORE G. SHEPHERD**

**T**he planet is unequivocally warming<sup>1</sup>, but the rate of warming varies in time and space. Although such variations are not unexpected, and do not challenge our grasp of the physics of global warming, understanding what drives them is crucial for predicting climate conditions at the regional scale. On page 465 of this issue, Horton *et al.*<sup>2</sup> use a new analytical method to argue that, over many regions and seasons, changes in the occurrence of atmospheric circulation patterns were a substantial driver of the observed changes in mid-latitude temperature extremes in the Northern Hemisphere over land for the 35 years to 2013. Such changes have increased not only the likelihood of summertime heatwaves in Europe, but also, since 1990, the

occurrence of low wintertime temperature extremes across central Asia<sup>3</sup>.

Identifying changes in climate extremes at the regional scale is hampered by the chaotic internal variability of the climate system, which adds a large degree of noise to the climate-change signal. Climate variability is manifested in changes in the occurrence of atmospheric circulation patterns, which can vary strongly from year to year, but can show trends over multidecadal periods too. Such variations are also reflected in many climate extremes<sup>4</sup>. For example, European heatwaves (Fig. 1) are more frequent when the summertime jet stream is 'blocked' from taking its usual path across the North Atlantic Ocean, because stagnant air over the continent can warm excessively through land–surface feedbacks<sup>5</sup>.

On top of this internal variability, climate



FRANK PERRY/AFP/GETTY

**Figure 1 | The River Loire in France during the 2003 European heatwave.** Horton *et al.*<sup>2</sup> report that this heatwave was part of a long-term trend towards more anticyclonic summertime circulation regimes over Europe.

change itself is expected to lead to changes in the occurrence of atmospheric circulation patterns, although there is as yet little consensus on what these changes will be<sup>1</sup>. Both internal variability and the circulation response to climate change can lead to regional changes in temperature extremes that are larger or smaller than — or even opposite to — those expected from global warming over extended time periods<sup>6</sup>.

To deal with the confounding effects of atmospheric circulation on changes in temperature extremes, Horton *et al.* applied a cluster analysis to geopotential height fields — which map the heights needed to reach a given pressure, taking into account variations in surface pressure and atmospheric density — at an altitude of about 5 kilometres, to identify robust changes in the occurrence of circulation regimes. Their approach restricts the analysis to the years after 1979, when observations from meteorological satellites improved the reliability of analysed changes in atmospheric circulation at mid-latitudes. But it avoids previously voiced concerns about using subjective measures of atmospheric circulation<sup>7</sup>. The authors then partitioned observed changes in temperature extremes into those associated with changes in the occurrence of circulation regimes (dynamic changes) and those that have no such association (thermodynamic changes). This kind of partitioning is a growing theme in climate science<sup>8</sup>.

Horton and colleagues find that most regions have experienced a clear increase in summertime high temperature extremes since 1979. For roughly half of those regions (eastern

North America, Europe and western Asia), the researchers identify a substantial contribution to the observed trends (one-third to one-half) from the increased occurrence of anticyclonic circulations, which are associated with atmospheric blocking. The implication is that the tendency for increased summertime heatwaves that can be expected from global warming has been substantially enhanced in these regions over the past 35 years by circulation changes that make heatwaves more likely.

Although global warming suggests that, in general, the likelihood of low temperature extremes should decrease, Horton *et al.* find an increase in wintertime low temperature extremes over central Asia since 1990 — the period over which the extent of Arctic sea ice has rapidly declined. They attribute this increase primarily to the increased occurrence of a circulation pattern associated with the transport of cold Arctic air into central Asia. Modelling studies<sup>9</sup> have suggested that such an altered circulation is attributable to declining sea-ice extent in the Barents and Kara seas. However, the authors find no evidence of increased wintertime low temperature extremes over North America, despite the attention given to claims of such an increase<sup>3</sup>.

One concern with this study is that it is based on 35-year (and, in some cases, only 24-year) time series, which are quite short by climate standards. Horton and co-workers performed an extensive analysis to determine the statistical significance of the trends, but this establishes only that discernible changes occurred over the periods in question. As the

authors note, the identified circulation trends may merely reflect multidecadal natural variability and cannot be associated definitively with climate change. This is clearly an important issue to resolve. Other climate extremes, such as droughts and flooding, may be even more susceptible to circulation changes than are temperature extremes.

A better observational record of past changes in atmospheric circulation, over the previous century and possibly over even longer timescales, would help to put the recent changes in context. Such records might be acquired by analysing data from ships' log-books or by developing climate proxies, for example. Improved prediction capabilities also require a better understanding of the mechanisms of climate variability. Finally, a better understanding of the physics of the expected atmospheric circulation response to climate change would provide testable hypotheses. In particular, can the most extreme predictions of circulation changes be constrained or excluded by the observational record? Unless progress is made on these fronts, we will continue to be surprised by deviations from the thermodynamic expectations of global warming. ■

**Theodore G. Shepherd** is in the Department of Meteorology, University of Reading, Reading RG6 6BB, UK.

e-mail: theodore.shepherd@reading.ac.uk

1. IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. *et al.*) (Cambridge Univ. Press, 2013).



- Horton, D. E. *et al. Nature* **522**, 465–469 (2015).
- Cohen, J. *et al. Nature Geosci.* **7**, 627–637 (2014).
- Screen, J. A. & Simmonds, I. *Nature Clim. Change* **4**, 704–709 (2014).
- Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C. & Vilà-Guerau de Arellano, J. *Nature Geosci.* **7**, 345–349 (2014).
- Deser, C., Phillips, A. S., Alexander, M. A. & Smoliak, B. V. J. *Clim.* **27**, 2271–2296 (2014).
- Barnes, E. A. *Geophys. Res. Lett.* **40**, 4734–4739 (2013).
- Shepherd, T. G. *Nature Geosci.* **7**, 703–708 (2014).
- Mori, M., Watanabe, M., Shiogama, H., Inoue, J. & Kimoto, M. *Nature Geosci.* **7**, 869–873 (2014).

## BIOCHEMISTRY

# Unexpected role for vitamin B2

An enzyme has been found that alters the molecular structure of vitamin B2, adding a fourth ring to its existing three-ring system. The product catalyses new types of chemistry in concert with certain other enzymes. [SEE LETTERS P.497 & P.502](#)

CATHERINE F. CLARKE  
& CHRISTOPHER M. ALLAN

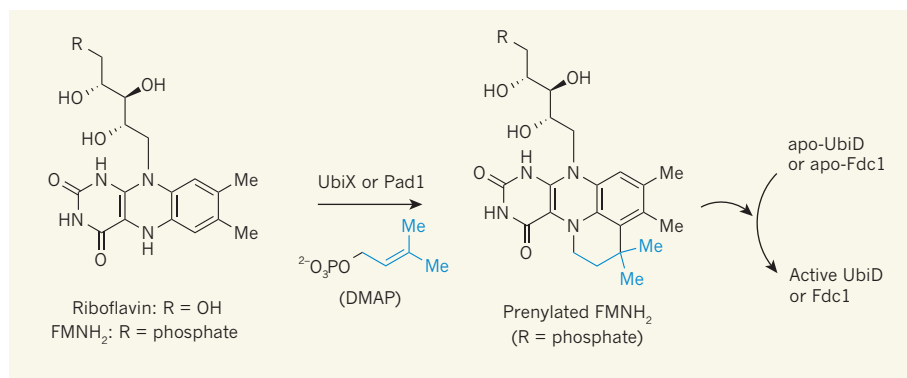
Many vitamins work by expanding the biochemical repertoire of enzyme-catalysed reactions. For example, riboflavin (vitamin B2) does this for a wide variety of flavoproteins, which catalyse many of the redox reactions in central metabolic pathways and act in the electron-transport chains of cells. Two papers in this issue, by White *et al.*<sup>1</sup> (page 502) and Payne *et al.*<sup>2</sup> (page 497), report a previously unknown cofactor derived from riboflavin. The findings solve the long-standing mystery of how a pair of bacterial enzymes — and their counterparts in yeast — catalyse crucial reactions known as decarboxylations, and further expand the repertoire of enzymatic reactions.

White *et al.* relate that the molecular structure of riboflavin, which contains three rings, can be modified by the addition of a prenyl group (a hydrocarbon group containing five

carbon atoms, also known as an isoprenyl group) to form a fourth ring (Fig. 1). The authors call the resulting compound prenylated FMNH<sub>2</sub>. They used high-resolution crystal structures, biochemical studies, spectroscopy and computational calculations to characterize prenylated FMNH<sub>2</sub> and to determine how it is employed as a cofactor in decarboxylation reactions — in which a carboxylate group (CO<sub>2</sub><sup>−</sup>) is removed in the form of carbon dioxide from a substrate.

Surprisingly, the enzyme that forms prenylated FMNH<sub>2</sub> is UbiX, a protein first found in the bacterium *Escherichia coli*. The authors show that UbiX works in tandem with another enzyme, called UbiD, to mediate the decarboxylation of an intermediate in the biosynthesis of coenzyme Q (see Fig. 1a of the paper<sup>1</sup>). Coenzyme Q is an essential lipid in the electron-transport chain and a potent cellular antioxidant<sup>3</sup>.

How exactly UbiX and UbiD cooperate



**Figure 1 | Prenylated riboflavin.** Flavin mononucleotide (FMNH<sub>2</sub>) is a common enzyme cofactor that is derived from riboflavin (vitamin B2). White *et al.*<sup>1</sup> show that the bacterial UbiX enzyme and the fungal Pad1 enzyme modify FMNH<sub>2</sub> in the presence of dimethylallyl monophosphate (DMAP), a precursor of prenyl groups, to generate a previously unknown cofactor, prenylated FMNH<sub>2</sub>. Payne *et al.*<sup>2</sup> show that prenylated FMNH<sub>2</sub> is an essential cofactor for the decarboxylase enzymes UbiD and Fdc1; apo-forms of UbiD and Fdc1 lack a bound cofactor and so are inactive. Me, methyl group.



## 50 Years Ago

As an introduction, I should like to touch on something which is not a unique feature of the subject of discourse. However, this is as good an excuse as any to look at the disturbing fact that communications between scientist and scientist are in a state of overgrowth or 'overpublication', while in contrast to this the writings ... addressed to the non-expert are, so some believe, somewhat neglected or 'underpublished'. Of course, one does not condemn the vastly increased potentialities (5–10 times that of 1945) for scientific printing space, but the sometimes indiscriminate multiplication of the printed words, formulae and illustrations leads to overburdening of libraries, to a nightmarish frustration of the research worker who never catches up with his reading, and willy-nilly to an encouragement of unfinished and mediocre material to be published.

From *Nature* 26 June 1965

## 100 Years Ago

All good nomenclature should be unambiguous, and, if possible, self-explanatory. The terms *masse volumique*, *volume massique*, and *stéradian* have both these desirable qualities; no one with a knowledge of physics and French could make any mistake as to the exact meaning of the first two, and the meaning of the third should be at once self-evident to anyone who knows the definition of a solid angle. I should not expect a chemist or a botanist to have anything but a hazy idea of the meaning of *puissance massique*, but even to an ordinary French engineer it should convey its meaning instantly. An expression of this kind, far from being an "eccentricity", is a triumph of nomenclature. It is possible to mould language by logic; it is the only way to mould language that shall be truly scientific.

From *Nature* 24 June 1915

to perform the decarboxylation reaction has been quite puzzling. Both enzymes are present in a wide array of prokaryotes (bacterial and archaeal microorganisms), and many other microbes, including fungi, rely on homologous enzymes — proteins that share ancestry with UbiD and UbiX — to decarboxylate various acids. White *et al.* show that prenylated FMNH<sub>2</sub> is first oxidized by Fdc1, a fungal homologue of UbiD, and then used to decarboxylate cinnamic acid (see Fig. 1a of the paper<sup>1</sup>).

So how does prenylated FMNH<sub>2</sub> form? White *et al.* reveal that when UbiX is supplied with dimethylallyl monophosphate (a precursor of prenyl groups), it prenylates riboflavin's three-ring system. This results in the formation of a fourth ring to generate prenylated FMNH<sub>2</sub> (Fig. 1). The authors show that this previously unknown cofactor enables Fdc1 to decarboxylate substrates *in vitro*, in the absence of UbiX. These results explain the findings of other studies<sup>4,5</sup> that showed that an unknown small molecule — but not UbiX itself — is required to activate the ability of UbiD (and Fdc1) to decarboxylate substrates.

Payne *et al.*<sup>2</sup> show that apo-Fdc1 (an apoenzyme is one without its bound cofactor) catalyses unusual biochemistry only when in the presence of prenylated FMNH<sub>2</sub>. They find that Fdc1 first oxidizes prenylated FMNH<sub>2</sub> to generate an 'iminium' form of the cofactor. This activates and prepares Fdc1 to decarboxylate a wide assortment of substrates known as  $\alpha,\beta$ -unsaturated aromatic carboxylic acids. Many of these derive from the microbial breakdown of lignin, the structural component of the secondary cell walls of plants<sup>6</sup>. The authors propose a mechanism for these decarboxylation reactions known as 1,3-dipolar cycloaddition (see Fig. 4d of the paper<sup>2</sup>). Although this mechanism is well known to organic chemists, its use by enzymes has until now been speculative<sup>7</sup>.

The findings raise several questions about the biological use of prenylated riboflavin cofactors. First, there is no obvious amino-acid sequence for binding prenylated FMNH<sub>2</sub>; UbiX and UbiD (or Fdc1) do not share similar amino-acid sequences or structures, and the interactions of these proteins with the prenyl group seem to be restricted to providing an appropriately shaped binding site. This complicates the identification of other enzymes that might make use of this cofactor.

It is also surprising that UbiX uses dimethylallyl monophosphate as a source of prenyl groups, rather than the commonly used source dimethylallyl diphosphate. It will be important to determine the metabolic origin of dimethylallyl monophosphate and to explore whether it is used in other enzyme reactions. Dimethylallyl diphosphate is generated either through a series of biochemical reactions known as the mevalonate pathway, or through the methylerythritol phosphate pathway. In

both pathways, dimethylallyl diphosphate is reversibly converted to another compound, isopentenyl diphosphate. A third, related pathway has recently been characterized<sup>8</sup> in the archaeon *Thermoplasma acidophilum*. This alternative route directly produces isopentenyl monophosphate — could it be that this compound interconverts with dimethylallyl monophosphate, in the same way that dimethylallyl diphosphate and isopentenyl diphosphate interconvert?

Can Pad1 (the fungal homologue of UbiX) and Fdc1 also mediate decarboxylation steps in the biosynthesis of coenzyme Q? UbiX from *E. coli* and Pad1 from yeast have been shown to perform the same function: a mutant form of *E. coli* in which the *ubiX* gene is deleted cannot synthesize coenzyme Q, but the synthesis is restored<sup>9</sup> if the mutant is engineered to express yeast Pad1. The two current papers now identify both Pad1 and UbiX as FMN prenyltransferases, enzymes that synthesize prenylated FMNH<sub>2</sub> as a diffusible small molecule.

But what about Fdc1 and UbiD? Do these enzymes both recognize 3-polyprenyl-4-hydroxybenzoic acid (the substrate that is decarboxylated in the biosynthesis of coenzyme Q)? Unfortunately, 3-polyprenyl-4-hydroxybenzoic acid is not commercially available, and so the authors of the current papers could not perform direct assays of Fdc1

or UbiD with this substrate. But the synthesis of coenzyme Q is not impaired when the genes that express Pad1 and Fdc1 are both deleted from two species of yeast<sup>9,10</sup>. The enzyme responsible for the decarboxylation step of coenzyme Q biosynthesis in eukaryotes (organisms that include fungi, plants and animals) therefore remains an open question. ■

**Catherine F. Clarke and Christopher M. Allan** are in the Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, USA.  
e-mail: cathy@chem.ucla.edu

1. White, M. D. *et al.* *Nature* **522**, 502–506 (2015).
2. Payne, K. A. P. *et al.* *Nature* **522**, 497–501 (2015).
3. Crane, F. L. *Mitochondrion* **7**, S2–S7 (2007).
4. Leppik, R. A., Young, I. G. & Gibson, F. *Biochim. Biophys. Acta* **436**, 800–810 (1976).
5. Lin, F., Ferguson, K. L., Boyer, D. R., Lin, X. N. & Marsh, E. N. *ACS Chem. Biol.* **10**, 1137–1144 (2015).
6. Weng, J. K. & Chapple, C. *New Phytol.* **187**, 273–285 (2010).
7. Krenske, E. H., Patel, A. & Houk, K. N. *J. Am. Chem. Soc.* **135**, 17638–17642 (2013).
8. Vinokur, J. M., Korman, T. P., Cao, Z. & Bowie, J. U. *Biochemistry* **53**, 4161–4168 (2015).
9. Gulmezian, M., Hyman, K. R., Marbois, B. N., Clarke, C. F. & Javor, G. T. *Arch. Biochem. Biophys.* **467**, 144–153 (2007).
10. Mukai, N., Masaki, K., Fujii, T., Kawamukai, M. & Iefuji, H. *J. Biosci. Bioeng.* **109**, 564–569 (2010).

This article was published online on 17 June 2015.

#### CELL METABOLISM

## Sugar for sight

**Retinitis pigmentosa causes the death of cone cells, leading to blindness. A factor secreted from rod cells, RdCVF, promotes cone survival in a mouse model of the disease. It now emerges that RdCVF works by increasing glucose uptake in cones.**

CONNIE CEPKO & CLAUDIO PUNZO

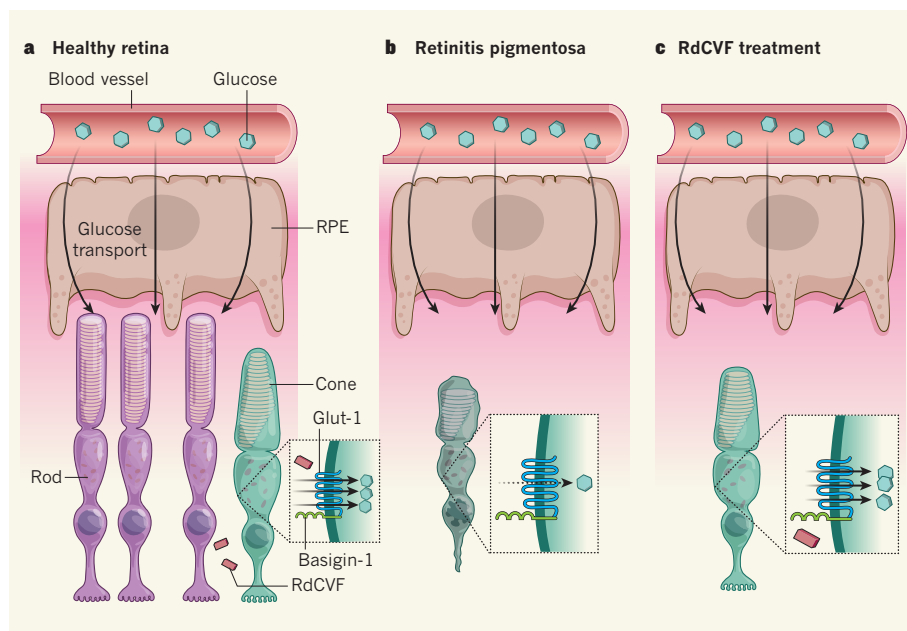
In daylight, light-sensing photoreceptor cells called cones endow us with rich, high-acuity colour vision. By contrast, their night-time counterparts, rod cells, are so light-sensitive that they enable us to see when walking in the woods under starlight. Sadly, a mutation in any one of more than 200 genes can cause diseases such as retinitis pigmentosa that bring about photoreceptor degeneration and lead to blindness<sup>1,2</sup>. Many of these disease-causing genes are rod-specific, affecting only night vision, but when rods die, the malfunction and death of cones soon follows. Writing in *Cell*, Ait-Ali *et al.*<sup>3</sup> provide evidence for an interaction between these two cell types that might explain why cones depend on rods for their survival, and that might eventually lead to a therapy for such diseases.

Retinitis pigmentosa affects 1 in 3,000 people<sup>2</sup>, and effective treatments are sorely needed. The

large number of disease-causing mutations means that these treatments should be gene-independent, focusing instead on targeting the biological pathways that cause the cones' death. The reasons for their death are unclear. However, there is evidence that several factors can contribute to cone death in retinitis pigmentosa. These include hyperoxia (an excess of oxygen), which causes oxidative damage by increasing the levels of free radicals<sup>4–6</sup>; a lack of energy<sup>7</sup>; and a lack of intermediates in the anabolic processes by which large molecules are constructed from smaller ones<sup>7</sup>.

In agreement with a role for such factors, hyperactivation of the protein complex mTOR1, which controls cell metabolism by balancing demand with supply, increases cone survival<sup>8</sup>. This protein complex probably acts by promoting the expression of genes that improve glucose uptake and use, raising levels of anabolic intermediates and of an anabolic cofactor molecule called NADPH. Adequate





**Figure 1 | Glucose uptake in cone cells.** **a**, Intermingled rod and cone photoreceptor cells in the outer portion of the retina receive glucose from blood vessels, through retinal pigmented epithelial (RPE) cells. Glucose uptake in photoreceptors is mediated by the transporter protein Glut-1, which interacts with another protein, Basigin-1. Ait-Ali *et al.*<sup>3</sup> demonstrate that a protein secreted from rods, called rod-derived cone viability factor (RdCVF), associates with Basigin-1 to improve glucose uptake and the production of energy from glucose by inducing the formation of a more active form of Glut-1. **b**, In diseases such as retinitis pigmentosa, in which rods die, cones seem to be unable to take up enough glucose to fuel their metabolism, and so also die. **c**, The authors show that this defect can be prevented by addition of RdCVF.

levels of NADPH are likely to be crucial to cone survival because, in addition to its role in anabolic processes, it is needed for pathways that detoxify free radicals in hyperoxic retinas. Injection of antioxidants<sup>4</sup> or viral-vector delivery of genes that fight oxidation<sup>6</sup> prolong cone survival in mouse models of retinitis pigmentosa, supporting the theory that oxidation is a cause of cone death.

Healthy photoreceptors are metabolically very active<sup>9</sup>, and so require high levels of glucose, which is delivered from the bloodstream through retinal pigmented epithelial cells. Because NADPH is produced by the oxidation of glucose, the demand for glucose in hyperoxic conditions is likely to be exceptionally high. A glucose transporter protein called Glut-1, located on the cell surface, mediates glucose uptake by photoreceptors, and evidence suggests<sup>7</sup> that a failure to take up sufficient glucose might contribute to cone death in retinitis pigmentosa. But there is a puzzling aspect to this model — glucose is delivered to the retina at a high rate and, after the death of rods, cones should have access to higher than normal levels of glucose. This suggests that there must be an added level of complexity underlying glucose uptake in cones.

Another factor that supports cone-cell survival<sup>10</sup> is a protein secreted from rods, called rod-derived cone viability factor (RdCVF)<sup>11</sup>, that may have antioxidant activity. The cones of mice lacking RdCVF are more susceptible to oxidative damage than those of controls, and

these mice show reduced photoreceptor activity with ageing<sup>12</sup>. Ait-Ali *et al.* therefore set out to explore the mechanism by which RdCVF promotes cone-cell survival. Using mass spectrometry, they identified a protein, called Basigin-1, that binds to RdCVF. Basigin-1 is found on the surface of cones and is known<sup>13</sup> to cause retinitis pigmentosa when mutated in mice. The authors also identified Glut-1 as a Basigin-1-binding protein. But a previous study showed that, contrary to what might have been expected, loss of Basigin-1 did not affect the expression of Glut-1 or, for the most part, its distribution in the retina<sup>14</sup>.

Ait-Ali and colleagues observed that addition of RdCVF increased glucose uptake, lactate release and ATP production in photoreceptor cells *in vitro* — three cellular responses suggesting that RdCVF increases metabolic flux. Furthermore, the authors found that a decrease in levels of Basigin-1 and Glut-1 eliminated the ability of RdCVF to promote photoreceptor survival. The authors propose that RdCVF, Basigin-1 and Glut-1 form a complex at the cell surface that increases glucose uptake (Fig. 1). However, the level of Glut-1 on the cell surface did not increase after RdCVF addition, leading the researchers to suggest that this complex instead acts to increase levels of the active form of Glut-1. Future work should test this model. It will also be of interest to study the potential antioxidant role of RdCVF, together with the formation and activity of this

three-protein complex, which might, as Ait-Ali *et al.* suggest, depend on a redox-sensitive interaction between RdCVF and Basigin-1.

A study published earlier this year showed that the delivery of RdCVF in mice with retinitis pigmentosa using an adeno-associated virus (AAV) prolonged cone survival and function<sup>15</sup>. AAV is a safe and effective vector that is used for ocular gene therapy in humans<sup>16</sup>, and this, together with Ait-Ali and colleagues' finding that Basigin-1 is expressed in human retinas, suggests that AAV-RdCVF might be an effective way to treat many types of photoreceptor disease. Owing to the large number of disease genes that cause blindness in humans, a treatment that could promote the survival of cones in a gene-independent manner would be a welcome prospect. ■

**Connie Cepko** is in the Departments of Genetics and of Ophthalmology, and at the Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. **Claudio Punzo** is in the Department of Ophthalmology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA.

e-mails: cepko@genetics.med.harvard.edu; claudio.punzo@umassmed.edu

1. <https://sph.uth.edu/retnet>
2. Hartong, D. T., Berson, E. L. & Dryja, T. P. *Lancet* **368**, 1795–1809 (2006).
3. Ait-Ali, N. *et al.* *Cell* **161**, 817–832 (2015).
4. Komeima, K., Rogers, B. S., Lu, L. & Campochiaro, P. A. *Proc. Natl Acad. Sci. USA* **103**, 11300–11305 (2006).
5. Yu, D.-Y., Cringle, S. J., Su, E.-N. & Yu, P. K. *Invest. Ophthalmol. Vis. Sci.* **41**, 3999–4006 (2000).
6. Xiong, W., MacColl Garfinkel, A. E., Li, Y., Benowitz, L. I. & Cepko, C. L. *J. Clin. Invest.* **125**, 1433–1445 (2015).
7. Punzo, C., Kornacker, K. & Cepko, C. L. *Nature Neurosci.* **12**, 44–52 (2009).
8. Venkatesh, A. *et al.* *J. Clin. Invest.* **125**, 1446–1458 (2015).
9. Graymore, C. *Br. J. Ophthalmol.* **43**, 34–39 (1959).
10. Léveillard, T. S. *et al.* *Nature Genet.* **36**, 755–759 (2004).
11. Holmgren, A. *Annu. Rev. Biochem.* **54**, 237–271 (1985).
12. Cronin, T. *et al.* *Cell Death Differ.* **17**, 1199–1210 (2010).
13. Hori, K. *et al.* *Invest. Ophthalmol. Vis. Sci.* **41**, 3128–3133 (2000).
14. Philp, N. J., Ochrietor, J. D., Rudoy, C., Muramatsu, T. & Linser, P. J. *Invest. Ophthalmol. Vis. Sci.* **44**, 1305–1311 (2003).
15. Byrne, L. C. *et al.* *J. Clin. Invest.* **125**, 105–116 (2015).
16. Trapani, I., Banfi, S., Simonelli, F., Surace, E. M. & Auricchio, A. *Hum. Gene Ther.* **26**, 193–200 (2015).

## ADDENDUM

The News & Views article 'Alzheimer's disease: From big data to mechanism' by Vivek Swarup and Daniel H. Geschwind (*Nature* **500**, 34–35; 2013) commented on the paper 'Integrative genomics identifies APOE ε4 effectors in Alzheimer's disease' by H. Rhinn *et al.* (*Nature* **500**, 45–50; 2013). This paper has now been retracted. For further information, see <http://dx.doi.org/10.1038/nature14591>

# The changing form of Antarctic biodiversity

Steven L. Chown<sup>1</sup>, Andrew Clarke<sup>2</sup>, Ceridwen I. Fraser<sup>3</sup>, S. Craig Cary<sup>4</sup>, Katherine L. Moon<sup>1,3</sup> & Melodie A. McGeoch<sup>1</sup>

**Antarctic biodiversity is much more extensive, ecologically diverse and biogeographically structured than previously thought. Understanding of how this diversity is distributed in marine and terrestrial systems, the mechanisms underlying its spatial variation, and the significance of the microbiota is growing rapidly. Broadly recognizable drivers of diversity variation include energy availability and historical refugia. The impacts of local human activities and global environmental change nonetheless pose challenges to the current and future understanding of Antarctic biodiversity. Life in the Antarctic and the Southern Ocean is surprisingly rich, and as much at risk from environmental change as it is elsewhere.**

The past two decades have witnessed a revolution in the approaches used to document patterns in and the functional significance of the diversity of life. Knowledge of the global distributions of organisms now extends to many groups. These include terrestrial plants and animals<sup>1</sup>, marine taxa<sup>2</sup>, components of the soil biota<sup>3</sup>, and, increasingly, microorganisms<sup>4</sup>. Empirical and theoretical research is verifying the range of mechanisms, including energy and nutrient availability, and historical contingency, that are responsible for these patterns<sup>1,2</sup>, and the role of diversity in maintaining ecosystem services. New approaches are also revealing rapid changes in diversity across most ecosystems, their often negative implications for the future of the planet, and our struggle to address the problem<sup>5</sup>.

In tandem, biodiversity research has increased across the Antarctic and Southern Ocean. The growing interest has been driven in part by the realization that any fundamental quest to understand life's diversity requires exploration of the polar regions<sup>6,7</sup>. But it has also been underpinned by an increasing appreciation of the particular challenges faced by Antarctic species and ecosystems. Climate change<sup>8</sup>, and economic activity in the form of fishing, tourism, and scientific research itself<sup>9,10</sup>, are all having impacts on Antarctic diversity. The intricacy of Southern Ocean food webs, best known for the region's iconic seals, penguins, whales and albatrosses, has been further unveiled, highlighting the complexity of the simultaneous impacts of changing climates, sea ice distribution and fisheries on this system<sup>11,12</sup>. New marine research has also exposed unexpectedly high benthic diversity, including in the deep sea<sup>13,14</sup>. In terrestrial areas, the diversity of higher plants and animals is relatively low (for example, just two flowering plant species), but the diversity of lichens, bryophytes, invertebrates, and the microbiota is substantial<sup>15</sup>. Local and regional spatial patterning in these groups is extensive<sup>16</sup>, with the mechanisms underlying these patterns both recognizable from broader ecological theory and unusual in several respects.

Here we explore recent advances in our understanding of the biodiversity of the Antarctic. Given the many conservation challenges it faces, we also consider how well the region is performing in relation to an aspirational set of global conservation benchmarks, the Aichi Targets<sup>5</sup>. Biodiversity conventionally encompasses variability across the entire biological hierarchy, from genes to ecosystems and their functioning. In the Antarctic, considerable recent progress has been made on all of these fronts<sup>17</sup>. In this review we focus primarily on ecological diversity,

its drivers, and changes that are being wrought by both local anthropogenic and global environmental change. We first bring together recent studies that have advanced our understanding of the diversity of marine and terrestrial groups, including the microbiota, and their evolution in the region. We then show how ecological and historical processes shape diversity and how they might change in the future. Finally, we assess Antarctic conservation in a global context, and highlight the biodiversity research challenges facing the region. Doing so is especially timely given the escalating political prominence of the Antarctic, due to debates about use of its resources, and the growing extent of science, fishing and tourism and their likely impacts.

## Marine richness, species flocks and diversification

The traditional view of Antarctic marine metazoan diversity has been of a species-poor fauna constrained by a harsh environment. Some groups are genuinely depauperate in the Southern Ocean, notably teleost fish, gastropods, bivalves and some crustaceans. Initial work on these groups largely shaped early views of the fauna as a whole. By the turn of the century, however, benthic diversity was thought to be high, with much of the fauna remaining undiscovered or undescribed<sup>18</sup>. Substantial sampling and systematic campaigns associated with the Census of Marine Life have since supported this idea<sup>19</sup>. More than 8,000 valid marine species are now known for the Southern Ocean, most of which are benthic<sup>14</sup>. Many of the new taxa have come from sampling of the deep Southern Ocean. For example, campaigns in the deep Weddell Sea recovered 674 isopod species, of which more than 80% were new to science<sup>13</sup>. The Southern Ocean is clearly not the universally depauperate region it was originally thought to be. Benthic sampling has also improved understanding of the evolution of unusual groups such as the acorn worms. Individuals collected along the West Antarctic Peninsula and in the Ross Sea live in secreted translucent tubes, making them similar to acorn worms of the Middle Cambrian, demonstrating conservation of this behaviour for 500 million years<sup>20</sup>.

Recent work has consolidated the view that strong latitudinal marine diversity gradients, with declines in species richness to the poles, are more general for the northern than southern hemisphere<sup>21</sup>. Indeed, for groups such as hexacorals (excluding scleractinians), pycnogonids and bryozoans, current knowledge suggests that diversity in the Southern Ocean is comparable with many temperate and non-reef tropical

<sup>1</sup>School of Biological Sciences, Monash University, Victoria 3800, Australia. <sup>2</sup>British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge, CB3 0ET, UK.

<sup>3</sup>Fenner School of Environment and Society, Australian National University, Canberra, ACT 2601, Australia. <sup>4</sup>International Centre for Terrestrial Antarctic Research, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand.



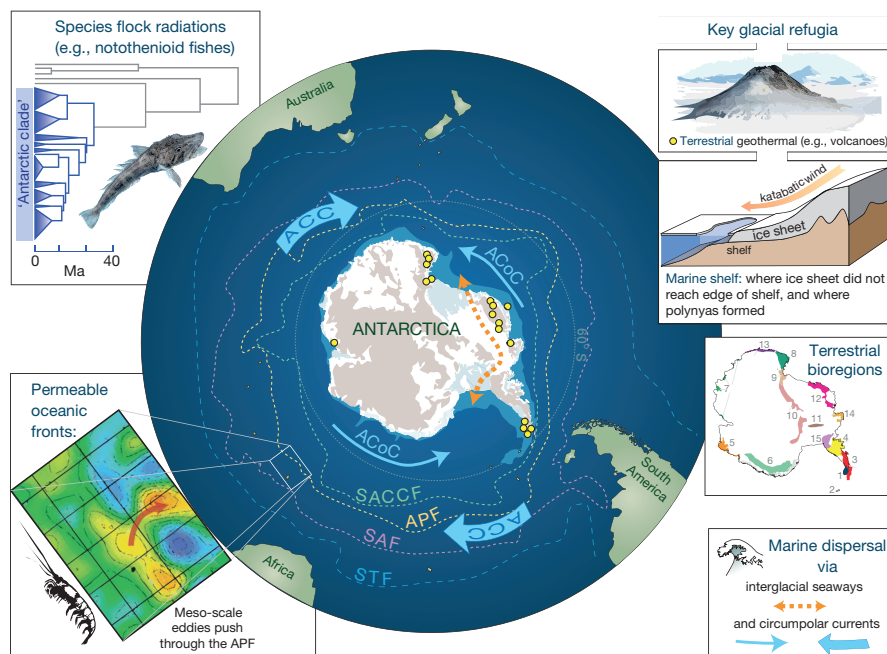
habitats<sup>14</sup>. In others, such as the petrels, albatrosses and their relatives, diversity patterns are atypical, with diversity peaks not in the tropics, nor at the highest southern latitudes, but rather in the productive areas between 35° S and 50° S<sup>22</sup>. The distinctiveness of the Southern Ocean fauna is also much higher than previously thought. Recent estimates suggest that 50 to 97% of Southern Ocean species in various groups such as sponges, tube worms, amphipods, molluscs, isopods, sea spiders and notothenioid fish are endemic<sup>14</sup>. The deep-sea fauna is also quite different to that of other regions<sup>13</sup>. For example, the hydrothermal vent faunas of the East Scotia Sea lack the polychaete worms, alvinocaridid shrimp and bathymodiolid mussels typical of other vent systems; instead these vents are characterized by peltospiroid gastropods, stalked barnacles, sea stars, and a previously unknown anomuran yeti crab species<sup>23</sup>. These differences suggest that the insights into metabolic functioning and the evolution of symbioses that have come from the exploration of deep-sea hydrothermal vents elsewhere<sup>24</sup> may be considerably extended through examination of Antarctic vent faunas.

Investigation of the high marine endemism has revealed how a complex set of earth system processes, including glaciation, differential diversification, and isolation, has interacted to shape the evolution of the southern biota<sup>25</sup>. In turn, the outcomes of this evolutionary interaction, such as the uniquely high prevalence in the region of marine invertebrate species lacking a planktonic larval phase<sup>26</sup>, have profound consequences for responses to ongoing global environmental change, and for forecasting them, owing to variation in dispersal capability.

One factor contributing to the rapid rise in estimates of diversity of the Southern Ocean has been the identification, typically using molecular approaches, of many cryptic species<sup>19</sup>. For example, within each of four *Orchomenella* amphipod species, several well-supported clades

have been found that probably represent reproductively isolated species<sup>27</sup>. Evidence of cryptic speciation is also present in Antarctic isopods<sup>28</sup>, echinoderms<sup>29</sup>, and sea slugs. In the latter, an original single species, *Doris kerguelensis*, is now thought to comprise a species flock, with diversification driven by predation<sup>30</sup>. Several other invertebrate groups, including echinoids, isopods and amphipods, may also qualify as marine species flocks<sup>31</sup>. Notothenioid fishes have long been recognized as a species flock, particularly those with antifreeze glycoproteins<sup>31</sup>, and the broad group includes several sub-clade radiations that have been linked to periods of climate change<sup>32</sup> (Fig. 1). Exploration of the way in which these species flocks have developed has provided much insight into evolutionary processes associated with rapid diversification, and those associated with survival in Antarctic waters<sup>33</sup>.

The cyclical waxing and waning of the ice sheets has been a major driver of range fragmentation and allopatric speciation in several marine benthic groups. Although much of the shelf and slope will have been inimical to life at glacial maxima<sup>34</sup>, and some shelf taxa undoubtedly retreated to the deep sea, palaeobiological and biogeographic evidence, coupled with recent molecular research, indicates that many species have persisted on the shelf throughout numerous glacial cycles<sup>35</sup>. These taxa could have survived glacial periods *in situ* in regions where the ice sheet did not extend to the edge of the continental shelf<sup>36</sup>, particularly where polynyas were formed by katabatic winds at shelf margins. Changes in sea level, together with partial collapses of the West Antarctic Ice Sheet between 1.1 million years and 125,000 years BP, may have allowed faunal exchange through seaways across the Antarctic continent, now closed by ice<sup>37</sup> (Fig. 1). These events have resulted in substantial biogeographic structuring<sup>38</sup>. The Southern Ocean is therefore not a relatively uniform region, but is biogeographically complex,



**Figure 1 | The Antarctic region is neither as isolated nor as depauperate in biodiversity as once thought.** Oceanic fronts can be permeable, with meso-scale eddies transporting pockets of water and associated organisms across them (inset figure lower left modified from ref. 109, with permission from Elsevier). Dispersal of marine species around Antarctica has been facilitated by the Antarctic Circumpolar Current (ACC) and Antarctic Coastal Current (ACoC), as well as by trans-Antarctic seaways during some interglacials<sup>37</sup>. Glacial cycles have driven diversification, with some species surviving within the Antarctic region during glacial periods—for example, in terrestrial refugia such as geothermal areas<sup>61</sup> or marine shelf refugia<sup>36</sup>—and undergoing allopatric divergence. Species flock radiations have been inferred for some

groups, such as Antarctic notothenioid fish which have diversified during periods of climate change (tree derived from ref. 32). The mean positions of the major oceanic fronts are shown: Southern Antarctic Circumpolar Current Front (SACCF), Antarctic Polar Front (APF), Subantarctic Front (SAF) and the Subtropical Front (STF). Locations of possible terrestrial geothermal refugia (volcanoes and heated rocks) are indicated by yellow dots<sup>61</sup>. Antarctic bedrock (>0 m) is roughly indicated in pale brown (derived from ref. 110), and the Antarctic continental shelf in blue (not shown for other continents). The diversity and position of the 15 distinct Antarctic Conservation Biogeographic regions<sup>44</sup> are indicated in the central panel on the right.

with considerably more local structuring than previously thought. Few taxa appear to be truly circumpolar, and in species initially believed to have broad (or even circumpolar) distributions subsequent phylogeographic work is revealing multiple species/divergent lineages with more restricted ranges<sup>27,29</sup>.

## Diversity in the terrestrial biota

Although flowering plant diversity is low in the Antarctic, other groups, such as the lichens (>200 species) and mosses (>100 species), are comparatively species-rich<sup>39</sup>. Some terrestrial invertebrate groups are also well represented, including the tardigrades, nematodes, springtails and mites<sup>40,41</sup>. The distribution patterns of these organisms are highly variable. Recent phylogeographic work shows that while some species appear truly circum-Antarctic in their distribution (for example, rotifers and nematodes<sup>41,42</sup>), others, such as springtails, are much more differentiated<sup>43</sup>. By coupling the growing availability of continent-wide biodiversity databases with ecoinformatic approaches, recent analyses have shown substantial spatial patterning in diversity across the continent, with fifteen distinct ecoregions now recognized on the continent itself, and another eight across the Southern Ocean islands<sup>44</sup>. Much diversity within species, at local spatial scales of hundreds of metres to hundreds of kilometres, is also being discovered through phylogeographic approaches, reflecting the effects of both older glacial history and more recent events<sup>7</sup>.

As with marine groups, substantial cryptic diversity has been unveiled in the terrestrial metazoan biota<sup>42,45</sup> and its diversity is probably much higher than is currently appreciated. Moreover, for most groups, including those that are considered well known, the full extent and structure of the continent's diversity remains partially understood because much of it has not been thoroughly surveyed<sup>46</sup>. Although the overall area of ice-free land is small, about 46,000 km<sup>2</sup> (about 1/3 the area of England) or less than 0.5% of the continent's surface area, a great deal of it is remote and difficult to access.

The bulk of Antarctica's terrestrial diversity lies in the less visible groups, represented by the microbiota. Metagenomics has substantially improved current understanding of both its structure and functioning<sup>6,47,48</sup>. Early, culture-dependent studies provided only a superficial glimpse of the resident terrestrial microbial diversity. Independent, comparative studies from around the continent, using next-generation sequencing technologies and metagenomic approaches, concur that terrestrial microbial community phylogenetic and taxonomic diversity are typically higher than once thought. Nonetheless, this diversity is lower than in many temperate soil systems, probably the result of selection imposed by the generally cold, dry, oligotrophic nature of the Antarctic soil system<sup>49–51</sup>, though in some areas soils are moist and enriched by the activities of seabirds and seals.

At the phylum level, Antarctic soils appear to be structured similarly to soils elsewhere, dominated by the *Acidobacteria*, *Proteobacteria*, *Actinobacteria* and *Bacteroidetes*. However, when examining individual 'species' composition, the picture is very different, suggesting selection for a highly adapted and specialized community<sup>52</sup>. Metagenomic approaches have revealed much higher community-level genetic plasticity than originally suggested for these extreme oligotrophic mineral soils, especially in autotrophic, heterotrophic, and diazotrophic lifestyles, along with significant representation of thermal, osmotic, dormancy and nutrient limitation stress-response pathways<sup>51,53</sup>. These metagenomic surveys are beginning to provide clues as to what might drive community assembly and interactions and ultimately their success. They have also identified the presence of antibiotic resistance genes and a prominence of phage signals<sup>48</sup>, suggesting that some level of cell competition and interaction may exist that could also play a critical role in structuring these communities where the role of grazers remains uncertain.

Terrestrial Antarctica also harbours a remarkable diversity of aquatic systems ranging from thousands of small ponds and lakes scattered across the terrestrial and ice shelf landscape to the recently accessed,

extensive network of subglacial lakes and streams. The lakes and ponds range from permanently ice covered to perennially ice free, freshwater to hypersaline, and highly stratified to permanently mixed, with much variation in geochemistry<sup>6</sup>. These features are all microbially dominated and, like Antarctic soils, generally support relatively low diversity communities when compared to similar temperate systems<sup>54</sup>. Metagenomic surveys suggest that this might not be the case for viruses, with Antarctic lakes supporting the highest diversities of viruses in any aquatic system yet examined<sup>15</sup>, emphasizing their functional and evolutionary significance<sup>55</sup>. Microbial mats, especially important systems in polar aquatic environments, also harbour high diversity, with members that include cosmopolitan and apparently endemic elements, both in mats on land and those on ice shelves<sup>56</sup>. Perhaps most noteworthy from an evolutionary perspective, the retrieval of samples from subglacial Lake Whillans has validated the assertion that life exists below the Antarctic ice sheets<sup>57</sup>. The lake supports a diverse, metabolically active, chemosynthetic community of Bacteria and Archaea, possibly relying on the oxidation of ammonia originating from microbial mineralization in the sediments laid down some 120,000 years ago.

## Drivers of diversity

Globally, the drivers of diversity have become increasingly well characterized, although controversy about their relative significance continues. Historical events such as glaciations and continental drift, along with dispersal have played major roles. Energy availability, in the form of primary productivity, is likewise important for determining variation in species richness<sup>1</sup>. Other factors influencing diversity include temperature, surface area, spatial and temporal heterogeneity in conditions, and interactions among organisms such as through predation and disease. Spatial scale and location play major roles in determining which of these factors are most significant. For Antarctic marine and terrestrial systems the same sets of drivers are often important.

In marine systems, richness varies with available energy, disturbance, and local heterogeneity, while climate cycles and long-term oceanic conditions have influenced which groups are diverse and which are less well represented<sup>7</sup>. Nonetheless, some of the drivers of diversity in the Antarctic are unusual, providing unique insights into the ongoing controversy. For example, spatial variation in the global richness of albatrosses, petrels and their relatives, which peak in diversity in the region, is positively related to the availability of large, contiguous areas of ocean, temperature and wind speed, all of which nonetheless represent forms of available energy<sup>22</sup>. Wind energy is especially interesting because it is rarely considered during investigations of the drivers of species richness, but can clearly be important, as studies of individual species of seabirds have demonstrated<sup>58</sup>. Moreover, winds are strengthening substantially in the region as a consequence of a southward shift and increase in intensity of the Westerlies, partly as a result of anthropogenic ozone depletion<sup>8</sup>. For species which rely on wind energy these changes should prove beneficial. Recent analyses of long-term life history and population data on wandering albatrosses support such an effect<sup>59</sup>. For breeding adults on the Crozet Islands (located at 46° S in the Southern Indian Ocean), rates of travel and flight speeds have increased and the duration of foraging trips has decreased over the past 20 years. Breeding success has consequently improved. Breeding birds have also gained about 1 kg in mass (individual birds weigh about 8–11 kg), which may represent an adaptive response to windier conditions.

In terrestrial Antarctic systems, as is the case for those elsewhere, the availability of energy and water are the two most important drivers of metazoan diversity. Across broader scales, moss, lichen and liverwort species richness is positively related to temperature<sup>16,60</sup>, though south of 72° S a suite of local microclimate factors including temperature, available water, light and shelter may be more important than energy alone. These gradients emphasize the importance of temperature in dictating the availability of water, though at local scales individual species distributions are typically also influenced by soil parameters, such as carbon and nitrogen availability<sup>7,42</sup>. Where geothermal energy warms the



ground, the availability of resources is higher both spatially and seasonally. Such consistent energy availability has played a major historical role in determining patterns of diversity across the continent. Analyses of a large database on the distribution of diversity in the region have revealed that for plants and fungi (mostly lichens) the presence of geothermal sites, such as volcanoes, explains much of the variation in diversity. This pattern extends beyond the immediate vicinity of these areas, suggesting that such sites have served as refugia for long periods, including during the Pleistocene glaciations<sup>61</sup>. Refugia such as these, as well as inland nunataks and coastal areas, have probably all had major roles in determining the distribution of many of the continent's terrestrial organisms<sup>62</sup>. The extent to which the importance of these drivers varies among different groups remains unknown, however, because of limited biodiversity surveys. Nonetheless, the importance of energy availability and refugia in terrestrial systems mean that the drivers of diversity for metazoans here have many parallels with those of the Antarctic marine environment.

Although most macrofauna and even some micro-eukaryotes show strong continental regionalization, much local structural heterogeneity in soil microbial communities has been revealed by regional and landscape-scale studies. Prokaryotic diversity variation, especially south of 70° S, may thus be driven predominantly by local microclimate and landscape history<sup>63,64</sup>. Soil surface ages and frequency of water availability establish severe gradients in pH, soil moisture and salinity that appear to be primary drivers for microbial community selection. Certain dominant components of the edaphic community, such as cyanobacteria, may be structured both by niche-related factors, such as in the case of communities living below and within rocks<sup>64</sup>, and in other circumstances by dispersal-related spatial filters (that is, wind processes) at a local scale. Recent studies examining the presence, composition, and local contributions of a possible continental-wide Aeolian microbial community suggest that this community is primarily of Antarctic origin with some contribution of more resilient taxa from off-continent sources<sup>65</sup>. The Antarctic aerosolized community appears to have little impact on the less optimal cold, dry, edaphic mineral soil communities, but makes a substantial contribution to the substantially more selective geothermal soils<sup>66</sup>.

For most lakes and ponds, geochemistry appears the primary driver for microbial community composition and structure, in keeping with the situation for many water bodies globally. Although mixing between nearby and disparate ponds and lakes probably occurs<sup>67,68</sup>, those with similar geochemistry (that is, conductivity and/or nutrients) support similar communities, at least at the higher taxonomic level, regardless of location<sup>69,70</sup>. This appears especially true for the more isolated bottom water of permanently stratified lakes, as evidenced by the high level of sequence divergence seen in certain groups<sup>71</sup>. Even some of the micro-eukaryotes appear to have very specific geochemically driven requirements, while others are considerably more cosmopolitan<sup>72</sup>.

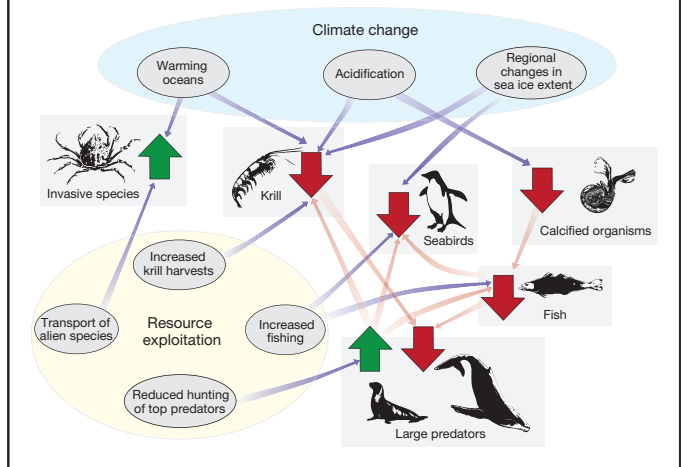
## Diminishing isolation

An enduring concept is that the Antarctic biota has been isolated since the breakup of Gondwana. Geologically, the terrestrial link between Antarctica and South America was severed in the Eocene, and deep-water flow of the Antarctic circumpolar current (ACC) was initiated in the Late Oligocene<sup>73</sup>. With the onset of the ACC came circumpolar thermal fronts such as the subtropical front (STF) and Antarctic polar front (APF) (Fig. 1), which define the boundary zones between Antarctic, sub-Antarctic and temperate waters. Steep temperature gradients and limited water exchange across these zones created dispersal barriers, exacerbated by strong circumpolar winds and currents that promote eastward dispersal, but hinder north-south movement<sup>74,75</sup>. Although the APF has been a formidable challenge to faunal exchange, it is not an absolute barrier<sup>76,77</sup>. Phylogeographic research now supports the idea of dispersal across the APF in a range of marine groups, such as molluscs<sup>78,79</sup>, echinoderms<sup>80,81</sup>, and crustaceans<sup>74</sup>. Many successful colonizations appear to have preceded Pliocene intensification of the ACC<sup>77</sup>,

## BOX 1

### Impacts of anthropogenic environmental change on Southern Ocean systems

Conceptualized impacts of local (resource exploitation, invasive species) and global (climate change, including increasing CO<sub>2</sub> concentrations which result in ocean acidification) environmental change drivers on Southern Ocean systems. Direct impacts are shown by light blue arrows and indirect impacts by light red ones. The large red and green arrows indicate declines or increases in populations, respectively. Human impacts commenced with the exploitation of whales and seals, which has now largely ceased. The main focus of exploitation is now fish<sup>10,98</sup> and krill, which not only has direct impacts on their populations, but can also affect other species which either depend on these organisms as a resource<sup>105</sup>, or are caught incidentally (such as some seabirds). Fishing outside the region also influences seabird populations as a consequence of incidental mortality. Climate change is having complex impacts through changes in sea ice distribution and temperature<sup>97,105</sup>. In the future, limited physiological tolerances and their capacity for change may have considerable effects both on vertebrates and invertebrates<sup>11,106</sup>. Ocean acidification is already affecting calcified organisms<sup>107</sup>, and may influence krill via impacts on eggs<sup>108</sup>. Although records of marine invasive species are uncommon, introductions, such as of crabs, may pose considerable threats to local communities<sup>88</sup>.



but occasional dispersal events into and out of Antarctic waters have occurred since that epoch<sup>75,82</sup>. Similar dispersal events across the APF, over the same timescales, have occurred in terrestrial bryophytes and invertebrates<sup>45,46</sup>. For important elements of the terrestrial microfauna, such as the tardigrades, rotifers and nematodes, information remains too sparse to reach general conclusions<sup>41</sup>, although long-term isolation in the continental Antarctic seems to be the case for some groups such as the mites and springtails<sup>62</sup>.

The rarity of natural dispersal events has, until recently, allowed Antarctica's biota to evolve and diversify in relative isolation<sup>83</sup>. Signals of natural dispersal, colonization and diversification for Antarctica are now at risk of being overwhelmed by impacts associated with changing climates and rapidly increasing human movement both into the region and between its distinct ecoregions. For example, 20% of tourist and 45% of scientific visitors to Antarctica carry plant propagules from species growing elsewhere, with more than 70,000 seeds estimated to be entering the region annually<sup>84</sup>. Many of these are from regions such as the Arctic and sub-Antarctic, where conditions are similar to more

**Table 1 | The relevance to the Antarctic of the 20 Aichi Targets under the Strategic Plan for Biodiversity 2011–2020 of the Conference of Parties to the Convention on Biological Diversity.**

| Aichi Target topic areas |   | Relevance of Target to the Antarctic  |
|--------------------------|---|---|
| 1*                       | Awareness of biodiversity values.   | While some Antarctic Treaty nations have coordinated programs promoting awareness of the value of the region, others do not. No coordinated regional program to do so exists.   |
| 2*                       | Biodiversity values integrated into planning and reporting.               | Biodiversity values are recognized in the Antarctic Treaty System through the CEP.  |
| 3*                       | Incentives for conservation and sustainable use.                          | Most applicable to nations fishing in the region, though the extent of such incentives is not widely known.   |
| 4                        | Impacts of natural resource use maintained within safe ecological limits. | No agricultural production takes place in the Antarctic. Resource extraction (fisheries) supplies consumers outside of the region.  |
| 5                        | Reduction in the rate of habitat loss.                                    | Changing climates are altering both marine and terrestrial systems.   |
| 6*                       | Sustainable fisheries and aquatic ecosystems.                             | A substantial fishery exists, overseen using an ecosystem approach by the CCAMLR. Much concern exists about the sustainability of the fishery <sup>10,98</sup> .  |
| 7                        | Sustainable production systems.   | No area is currently under production.  |
| 8*                       | Pollution reduction.  | Point source pollution is limited with much effort to reduce it.  |
| 9*                       | Prevention and management of biological invasions.                        | Non-indigenous species are a major concern for the region, and are a priority for the CEP. Activity to control species and pathways is inconsistent and uncoordinated.  |
| 10*                      | Pressures on marine ecosystems impacted by climate change.                | Ocean acidification is a substantial threat to Southern Ocean ecosystems <sup>97,107,108</sup> .  |
| 11*                      | Targets for protected areas.  | Only 1.5% of terrestrial biodiversity area is formally protected. Management is inadequate <sup>93</sup> . The designation of MPAs is controversial.  |
| 12*                      | Improvement in species conservation status.                               | Knowledge of the conservation status of taxa is poor, other than for most birds and mammals.  |
| 13                       | Safeguarding genetic diversity of valuable species.                       | Largely not relevant, although bioprospecting is increasing <sup>100</sup> . Southern Ocean ecosystems are important for whales, which have cultural value.   |
| 14*                      | Restoration of ecosystem services.  | Antarctic ecosystems provide several services essential to planetary stability, including maintenance of land ice to prevent sea level rise, and climate regulation.  |
| 15                       | Contribution to climate change mitigation.                                | Not relevant to terrestrial systems. Its likely relevance is through marine system contributions to CO <sub>2</sub> sequestration.  |
| 16*                      | Fair and equitable access and benefit sharing.                            | Bioprospecting takes place widely in the region, with no agreement on access and benefit sharing <sup>100</sup> .   |
| 17*                      | Up to date biodiversity strategy and action plans.                        | Most of the nations with activities in the region are party to the CBD and have national biodiversity strategy and action plans. The Antarctic region is seldom considered in them. No comparable regional plan exists. |
| 18                       | Respect for traditional knowledge.  | No indigenous or local communities in the region are recognized by the Antarctic Treaty System.   |
| 19*                      | Improved knowledge of biodiversity.                                       | Scientific activity is supported by many national programmes, communicated inter alia via the Scientific Committee on Antarctic Research.   |
| 20*                      | Mobilization of financial resources.                                      | No mechanisms exist for mobilizing financial resources in support of a strategic plan for biodiversity in the region.   |

Targets with an asterisk are most relevant to the region.

benign parts of the continent, making establishment in Antarctica more likely. Some of these species have already established themselves at several locations, such as annual bluegrass, *Poa annua*, which is a superior competitor to the only two indigenous vascular plant species found along the Antarctic Peninsula<sup>85</sup>. Newly arrived terrestrial species continue to be documented<sup>86</sup>, and climate change is expected to increase the likelihood of establishment and further impact by non-indigenous species.

Some parts of the Antarctic marine system may be similarly affected, especially shelf environments. Ship traffic to the Antarctic carries a diverse assemblage of non-indigenous species as fouling<sup>87</sup>, and such traffic is heavy to the Antarctic Peninsula. Here, one widely-discussed risk is the spread of crabs with the potential for substantial changes to the benthic system<sup>88</sup> (Box 1). What the risks are for invasion of other areas of the Antarctic by marine groups is not clear, with understanding for the broader region poorly developed compared with terrestrial systems, stimulating controversy about the likely significance of marine invasions<sup>89</sup>.

### Antarctica and the strategic plan for biodiversity

Knowledge of the negative consequences of global change for the increasingly apparent scope and novelty of Antarctic biodiversity is growing (Box 1). A useful perspective on how well conservation of the region is performing by global standards comes from the Convention on Biological Diversity's (CBD) Aichi Targets of the Strategic Plan for Biodiversity 2011–2020<sup>5</sup>.

The Antarctic region is largely an international space governed predominantly through the Antarctic Treaty (which applies to all areas south of 60° S), to which 52 countries are party. Environmental matters are the main concern of its Committee for Environmental Protection (CEP) and the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR)<sup>90</sup>. In consequence, the region is typically not subject to many global agreements such as the CBD. Nonetheless, many of the Antarctic Treaty Parties are also signatory to these international agreements, including the CBD, and are actively developing plans to achieve the Aichi Targets<sup>91</sup>.

The Aichi Targets encompass 20 aspirational goals for global biodiversity conservation. Not all of them are immediately relevant to the Antarctic. For example, Targets 7, 13 and 18 are not germane because the activities and species they cover, and indigenous human populations, are typically absent from the region. Many of the targets could, however, be applied to the Antarctic (Table 1). Of these, only performance against Target 11, concerning protected areas, has been explicitly assessed<sup>92</sup>. Just 1.5% of the continent's ice-free area is formally designated as specifically protected, and 5 of its 15 ecoregions have no protection. Antarctic terrestrial biodiversity protection thus falls short of the 17% area coverage proposed in Target 11, and by some accounts management is insufficiently effective too<sup>93</sup>. Similar assessments have not been made for marine protected areas (MPAs), but only one exists in the Treaty Area. Together with recent controversy about the inability of the parties to CCAMLR to agree on Antarctic MPAs, this situation suggests



that for marine systems, area protection is poorer than for the terrestrial Antarctic.

For the remainder of the relevant Aichi Targets, the available evidence suggests that were assessments to be made, the region might fare poorly in all except perhaps Targets 8 and 9 (Table 1). The reduction of pollution to levels that are not detrimental to ecosystem function and biodiversity is the subject of Target 8. Much progress has been made in addressing pollution, though in some areas substantial problems still exist<sup>94</sup>. Target 9 specifies that by 2020, invasive alien species and pathways of introduction should be identified and prioritized, priority species controlled or eradicated, and measures put in place to prevent the introduction and establishment of such species. The Antarctic Treaty Parties have taken meaningful, evidence-based steps to identify pathways of invasion and priority species, and to prevent introduction and establishment, based on the advice of the CEP<sup>95</sup>.

## Outlook and challenges

The importance of science in and about the Antarctic region has recently been re-emphasized by the identification of 80 key questions that are on the scientific horizon for the region<sup>17</sup>. Many of these questions illustrate the major role science will play in mitigating the consequences of environmental impacts in the Antarctic and in providing the information required for evidence-based policy to do so both locally and globally. Perhaps more importantly from a fundamental science perspective, they also recognize that human activities in the Antarctic and elsewhere may be compromising current approaches and future ability to understand biodiversity and ecosystem functioning in the region.

From this perspective and from what is now known about biodiversity in the region, several challenges remain. First among these is inadequate knowledge of terrestrial diversity in many areas, for many taxa, and the absence of any coordinated program to rectify the situation. In the face of substantial activity across the region, and the clear indication of regional diversification in many groups, signals of phylogenetic and functional distinctiveness, especially in the microbiota, may be substantially compromised. Risks of introduction of non-indigenous taxa are not only from outside the continent, but also from regions that are significantly differentiated on the continent, at both continent-wide and local scales<sup>93</sup>. Comprehensive understanding of how diversity patterns and the mechanisms underlying them vary among major groups, and among major areas (such as above and below about 70° S), is yet to be developed, restricting comparisons with elsewhere and with marine systems. Biotic homogenization as a consequence of both scientific and non-scientific activities would certainly obscure understanding, though to what extent this might happen is itself unknown.

The importance of limiting local impacts from human activities has been further emphasized by recent work showing that responses to climate change, and in particular to pulse events, may be rapid<sup>96</sup>. Increases in water availability and carbon as a result of warmer, wetter summers, and progressively warming regional climates as seen in the Peninsula<sup>8</sup>, could radically alter biodiversity in the continent's ice-free areas on a near-immediate time scale, such as by increasing growth rates and activity periods, so increasing abundance and richness. Understanding the likelihood of such change and documenting it presupposes a limited influence of local disturbance.

A similar situation is true of marine systems (Box 1), where understanding of the functioning and dynamics of large, relatively undisturbed ecosystems remains an important goal of ecology. Until recently, this goal may well have been possible for the Antarctic given low impacts on the Ross Sea system, acknowledging the roles played by whales and seals that were removed in large numbers in the 19th and 20th centuries<sup>97</sup>. Now, that possibility is under threat. Changes in stratospheric ozone and greenhouse gases are driving substantial alterations to sea ice concentrations and seasonality, and ocean pH in the system<sup>97</sup>, and little progress has been made in securing the area from the local, and considerable, impacts of fishing<sup>98</sup>. Establishing a substantial MPA for

the Ross Sea is not only important for conservation, but also for the scientific value of retaining a large marine system that is not being simultaneously affected by both global anthropogenic drivers and local ones. Few such other marine systems exist. More generally, the situation reflects the absence of substantial national collaboration to achieve region-wide conservation objectives. No commitment to anything equivalent to the Aichi Targets exists for the region, and the Antarctic Treaty System remains underfunded, with new activities typically not reaching consensus because of budgetary concerns. In the face of growing challenges presented by fishing and interest in the Antarctic's other resources<sup>99,100</sup>, the situation is especially concerning.

Rapid developments in a range of areas nonetheless suggest that, from a science perspective, many of the challenges can be overcome. These include new DNA archival approaches that enable samples collected from rare visits to remote sites to be stored and re-investigated repeatedly<sup>101</sup>, satellite remote sensing that enables total population counts<sup>102</sup> and identification of areas with much metazoan diversity<sup>103</sup>, and new analytical tools to enable comparisons among multiple groups and regions<sup>104</sup>. Approaches that limit the influence of scientific activities on the very systems being investigated are also developing rapidly<sup>57</sup>. The application of these developments will mean a larger transformation in understanding of Antarctic biodiversity over the next 20 years than its change over the last century. In so doing investigations of a globally unusual and important system will continue to contribute to general understanding of the planet's diversity and ecosystems, and at the same time provide further evidence to support its stringent conservation.

Received 23 February; accepted 24 April 2015.

- Gaston, K. J. Global patterns in biodiversity. *Nature* **405**, 220–227 (2000).
  - Belanger, C. L. *et al.* Global environmental predictors of benthic marine biogeographic structure. *Proc. Natl Acad. Sci. USA* **109**, 14046–14051 (2012).
  - Tedersoo, L. *et al.* Global diversity and geography of soil fungi. *Science* **346**, 1256688 (2014).
  - Barberán, A., Casamayor, E. O. & Fierer, N. The microbial contribution to macroecology. *Front. Microbiol.* **5**, 203 (2014).
  - Tittensor, D. P. *et al.* A mid-term analysis of progress toward international biodiversity targets. *Science* **346**, 241–244 (2014).
  - Wilkins, D. *et al.* Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol. Rev.* **37**, 303–335 (2013).
  - Convey, P. *et al.* The spatial structure of Antarctic biodiversity. *Ecol. Monogr.* **84**, 203–244 (2014).
  - Turner, J. *et al.* Antarctic climate change and the environment: an update. *Polar Rec. (Gr. Brit.)* **50**, 237–259 (2014).
  - Tin, T. *et al.* Impacts of local human activities on the Antarctic environment. *Antarct. Sci.* **21**, 3–33 (2009).
  - Ainley, D. G. & Pauly, D. Fishing down the food web of the Antarctic continental shelf and slope. *Polar Rec. (Gr. Brit.)* **50**, 92–107 (2014).
  - Constable, A. J. *et al.* Climate change and Southern Ocean ecosystems I: how changes in physical habitats directly affect marine biota. *Glob. Change Biol.* **20**, 3004–3025 (2014).
  - Saba, G. K. *et al.* Winter and spring controls on the summer food web of the coastal West Antarctic Peninsula. *Nature Commun.* **5**, 4318 (2014).
  - Brandt, A. *et al.* First insights into the biodiversity and biogeography of the Southern Ocean deep sea. *Nature* **447**, 307–311 (2007).
- This study of benthic diversity challenged the notion that deep-sea diversity is depressed in the Southern Ocean, with its findings borne out by recent comprehensive surveys.**
- De Broyer, C. *et al.* *Biogeographic Atlas of the Southern Ocean* (Scientific Committee on Antarctic Research, 2014).
  - López-Bueno, A. *et al.* High diversity of the viral community from an Antarctic lake. *Science* **326**, 858–861 (2009).
- This study showed that an Antarctic lake viral community has high genetic richness distributed across the highest number of viral families found in aquatic viral genomes, with a substantial proportion of sequences related to eukaryotic viruses, unlike the situation for other aquatic viromes.**
- Casanovas, P., Lynch, H. J. & Fagan, W. F. Multi-scale patterns of moss and lichen richness on the Antarctic Peninsula. *Ecography* **36**, 209–219 (2013).
  - Kennicutt, M. C. II *et al.* Six priorities for Antarctic science. *Nature* **512**, 23–25 (2014).
  - Janosik, A. M. & Halanich, K. M. Unrecognized Antarctic biodiversity: a case study of the genus *Odontaster* (Odontasteridae; Asteroidea). *Integr. Comp. Biol.* **50**, 981–992 (2010).
  - Kaiser, S. *et al.* Patterns, processes and vulnerability of Southern Ocean benthos: a decadal leap in knowledge and understanding. *Mar. Biol.* **160**, 2295–2317 (2013).

20. Halanych, K. M., Cannon, J. T., Mahon, A. R., Swalla, B. J. & Smith, C. R. Modern Antarctic acorn worms form tubes. *Nature Commun.* **4**, 2738 (2013).
21. Clarke, A. in *Marine Macroecology* (eds Witman, J. D. & Roy, K.) 250–278 (University of Chicago Press, 2009).
22. Davies, R. G., Irlach, U. M., Chown, S. L. & Gaston, K. J. Ambient, productive and wind energy, and ocean extent predict global species richness of procariiform seabirds. *Glob. Ecol. Biogeogr.* **19**, 98–110 (2010).
23. Rogers, A. D. *et al.* The discovery of new deep-sea hydrothermal vent communities in the Southern Ocean and implications for biogeography. *PLoS Biol.* **10**, e1001234 (2012).
- This study showed that the fauna of deep-sea hydrothermal vents on the East Scotia Ridge in the Southern Ocean is wholly different to vent faunas elsewhere, demonstrating that Antarctic endemism extends to these faunas.**
24. Petersen, J. M. *et al.* Hydrogen is an energy source for hydrothermal vent symbioses. *Nature* **476**, 176–180 (2011).
25. Crame, J. A. Early Cenozoic differentiation of polar marine faunas. *PLoS ONE* **8**, e54139 (2013).
26. Marshall, D. J., Krug, P. J., Kupriyana, E. K., Byrne, M. & Emlet, R. B. The biogeography of marine invertebrate life histories. *Annu. Rev. Ecol. Syst.* **43**, 97–114 (2012).
27. Havermans, C., Nagy, Z. T., Sonet, G., De Broyer, C. & Martin, P. DNA barcoding reveals new insights into the diversity of Antarctic species of *Orchomene sensu lato* (Crustacea: Amphipoda: Lysianassoidea). *Deep Sea Res. Part II Top. Stud. Oceanogr.* **58**, 230–241 (2011).
28. Raupach, M. J., Malyutina, M., Brandt, A. & Wägele, J.-W. Molecular data reveal a highly diverse species flock within the munnopsoid deep-sea isopod *Betamorpho fusiformis* (Barnard, 1920) (Crustacea: Isopoda: Asellota) in the Southern Ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **54**, 1820–1830 (2007).
29. Wilson, N. G., Hunter, R. L., Lockhart, S. J. & Halanych, K. M. Multiple lineages and absence of panmixia in the “circumpolar” crinoid *Promachocrinus kerguelensis* from the Atlantic sector of Antarctica. *Mar. Biol.* **152**, 895–904 (2007).
- This broad-scale study challenges the generalization that many Antarctic species have circumpolar distributions, instead suggesting that much unrecognized diversity and geographic structure exists in the Antarctic biota.**
30. Wilson, N. G., Maschek, J. A. & Baker, B. J. A species flock driven by predation? Secondary metabolites support diversification of slugs in Antarctica. *PLoS ONE* **8**, e80277 (2013).
31. Lecointre, G. *et al.* Is the species flock concept operational? The Antarctic shelf case. *PLoS ONE* **8**, e68787 (2013).
32. Near, T. J. *et al.* Ancient climate change, antifreeze, and the evolutionary diversification of Antarctic fishes. *Proc. Natl Acad. Sci. USA* **109**, 3434–3439 (2012).
- This study showed that although antifreeze was acquired early in the evolution of notothenioid fishes in Antarctica, the main burst of diversification was much more recent, probably during the Late Miocene cooling.**
33. Czikó, P. A., DeVries, A. L., Evans, C. W. & Cheng, C. H. C. Antifreeze protein-induced superheating of ice inside Antarctic notothenioid fishes inhibits melting during summer warming. *Proc. Natl Acad. Sci. USA* **111**, 14583–14588 (2014).
34. Thatje, S., Hillenbrand, C.-D., Mackensen, A. & Larter, R. Life hung by a thread: endurance of Antarctic fauna in glacial periods. *Ecology* **89**, 682–692 (2008).
35. Rogers, A. D. in *Antarctic Ecosystems. An Extreme Environment in a Changing World* (eds Rogers, A. D., Johnston, N. M., Murphy, E. J. & Clarke, A.) 417–467 (Wiley-Blackwell, 2012).
36. Allcock, A. L. & Strugnell, J. M. Southern Ocean diversity: new paradigms from molecular ecology. *Trends Ecol. Evol.* **27**, 520–528 (2012).
37. Barnes, D. K. A. & Hillenbrand, C.-D. Faunal evidence for a late Quaternary trans-Antarctic seaway. *Glob. Change Biol.* **16**, 3297–3303 (2010).
- This ecological study showed striking similarities in bryozoan assemblages in the Weddell and Ross Seas, supporting the hypothesis that partial collapse of the West Antarctic Ice Sheet during Pleistocene interglacials created a trans-Antarctic seaway.**
38. Pierrat, B., Saucède, T., Brayard, A., David, B. & Crame, A. Comparative biogeography of echinoids, bivalves and gastropods from the Southern Ocean. *J. Biogeogr.* **40**, 1374–1385 (2013).
39. Peat, H. J., Clarke, A. & Convey, P. Diversity and biogeography of the Antarctic flora. *J. Biogeogr.* **34**, 132–146 (2007).
40. Stevens, M. I. & Hogg, I. D. in *Trends in Antarctic Terrestrial and Limnetic Ecosystems* (eds Bergstrom, D. M., Convey, P. & Huiskes, A. H. L.) 177–192 (Springer, 2006).
41. Velasco-Castrillón, A., Gibson, J. A. E. & Stevens, M. I. A review of current Antarctic limno-terrestrial microfauna. *Polar Biol.* **37**, 1517–1531 (2014).
42. Velasco-Castrillón, A. & Stevens, M. I. Morphological and molecular diversity at a regional scale: a step closer to understanding Antarctic nematode biogeography. *Soil Biol. Biochem.* **70**, 272–284 (2014).
43. Torricelli, G. *et al.* High divergence across the whole mitochondrial genome in the “pan-Antarctic” springtail *Friesia grisea*: evidence for cryptic species? *Gene* **449**, 30–40 (2010).
44. Terauds, A. *et al.* Conservation biogeography of the Antarctic. *Divers. Distrib.* **18**, 726–741 (2012).
45. Pisa, S. *et al.* The cosmopolitan moss *Bryum argenteum* in Antarctica: recent colonisation or *in situ* survival? *Polar Biol.* **37**, 1469–1477 (2014).
46. McGaughan, A., Stevens, M. I., Hogg, I. D. & Carapelli, A. Extreme glacial legacies: a synthesis of the Antarctic springtail phylogeographic record. *Insects* **2**, 62–82 (2011).
47. Vyverman, W. *et al.* Evidence for widespread endemism among Antarctic microorganisms. *Polar Sci.* **4**, 103–113 (2010).
48. Zablocki, O. *et al.* High-level diversity of tailed phages, eukaryote-associated viruses, and viroplasm-like elements in the metaviromes of Antarctic soils. *Appl. Environ. Microbiol.* **80**, 6888–6897 (2014).
49. Yergeau, E. *et al.* Size and structure of bacterial, fungal and nematode communities along an Antarctic environmental gradient. *FEMS Microbiol. Ecol.* **59**, 436–451 (2007).
50. Cary, S. C., McDonald, I. R., Barrett, J. E. & Cowan, D. A. On the rocks: the microbiology of Antarctic Dry Valley soils. *Nature Rev. Microbiol.* **8**, 129–138 (2010).
51. Fierer, N. *et al.* Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl Acad. Sci. USA* **109**, 21390–21395 (2012).
52. Lee, C. K., Barbier, B. A., Bottos, E. M., McDonald, I. R. & Cary, S. C. The inter-valley soil comparative survey: the ecology of Dry Valley edaphic microbial communities. *ISME J.* **6**, 1046–1057 (2012).
53. Chan, Y., Van Nostrand, J. D., Zhou, J., Pointing, S. B. & Farrell, R. L. Functional ecology of an Antarctic Dry Valley. *Proc. Natl Acad. Sci. USA* **110**, 8990–8995 (2013).
- This study showed, using a metagenomic approach, significant plasticity in autotrophic, diazotrophic and heterotrophic strategies which support microbial communities in the Antarctic Dry Valleys.**
54. Laybourn-Parry, J. & Pearce, D. A. The biodiversity and ecology of Antarctic lakes: models for evolution. *Phil. Trans. R. Soc. B* **362**, 2273–2289 (2007).
55. Anesio, A. M. & Bellas, C. M. Are low temperature habitats hot spots of microbial evolution driven by viruses? *Trends Microbiol.* **19**, 52–57 (2011).
56. Varin, T., Lovejoy, C., Jungblut, A. D., Vincent, W. F. & Corbeil, J. Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. *Appl. Environ. Microbiol.* **78**, 549–559 (2012).
57. Christner, B. C. *et al.* A microbial ecosystem beneath the West Antarctic ice sheet. *Nature* **512**, 310–313 (2014).
- This study showed that subglacial Lake Whillans, which lies below 800 m of ice, has a diverse, chemosynthetically driven assemblage of Bacteria and Archaea, thus verifying the existence of deep, subglacial life.**
58. Pennycuik, C. J. in *Comparative Physiology: Life in Water and on Land*, Vol. 9 (eds Dejours, P., Bolis, L., Taylor, C. R. & Weibel, E. R.) 371–386 (Liviana Press, 1987).
59. Weimerskirch, H., Louzao, M., de Grissac, S. & Delord, K. Changes in wind pattern alter albatross distribution and life-history traits. *Science* **335**, 211–214 (2012).
60. Green, T. G. A., Sancho, L. G., Pintado, A. & Schroeter, B. Functional and spatial pressures on terrestrial vegetation in Antarctica forced by global warming. *Polar Biol.* **34**, 1643–1656 (2011).
61. Fraser, C. I., Terauds, A., Smellie, J., Convey, P. & Chown, S. L. Geothermal activity helps life survive glacial cycles. *Proc. Natl Acad. Sci. USA* **111**, 5634–5639 (2014).
62. Hawes, T. C. Antarctica’s geological arks of life. *J. Biogeogr.* **42**, 207–208 (2015).
63. Barrett, J. E. *et al.* Co-variation in soil biodiversity and biogeochemistry in northern and southern Victoria Land, Antarctica. *Antarct. Sci.* **18**, 535–548 (2006).
64. Pointing, S. B. *et al.* Highly specialized microbial diversity in hyper-arid polar desert. *Proc. Natl Acad. Sci. USA* **106**, 19964–19969 (2009).
- This study showed that considerable microbial diversity exists as four distinct communities, including three lithic ones, in the hyper-arid Antarctic Dry Valleys.**
65. Pearce, D. A. *et al.* Microorganisms in the atmosphere over Antarctica. *FEMS Microbiol. Ecol.* **69**, 143–157 (2009).
66. Herbold, C. W., Lee, C. K., McDonald, I. R. & Cary, S. C. Evidence of global-scale aeolian dispersal and endemism in isolated geothermal microbial communities of Antarctica. *Nature Commun.* **5**, 3875 (2014).
67. Gordon, D. A., Priscu, J. & Giovannoni, S. Origin and phylogeny of microbes living in permanent Antarctic lake ice. *Microb. Ecol.* **39**, 197–202 (2000).
68. Archer, S. D., McDonald, I. R., Herbold, C. W. & Cary, S. C. Characterisation of bacterioplankton communities in the meltwater ponds of Bratina Island, Victoria Land, Antarctica. *FEMS Microbiol. Ecol.* **89**, 451–464 (2014).
69. Bowman, J. P., McCammon, S. A., Rea, S. M. & McMeekin, T. A. The microbial composition of three limnologically disparate hypersaline Antarctic lakes. *FEMS Microbiol. Lett.* **183**, 81–88 (2000).
70. Villaseca, J. A. *et al.* A close link between bacterial community composition and environmental heterogeneity in maritime Antarctic lakes. *Int. Microbiol.* **13**, 67–77 (2010).
71. Lauro, F. M. *et al.* An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J.* **5**, 879–895 (2011).
72. Bielewicz, S. *et al.* Protist diversity in a permanently ice-covered Antarctic lake during the polar night transition. *ISME J.* **5**, 1559–1564 (2011).
73. Lefebvre, V., Donnadieu, Y., Sepulchre, P., Swingedouw, D. & Zhang, Z.-S. Deciphering the role of southern gateways and carbon dioxide on the onset of the Antarctic Circumpolar Current. *Paleoceanography* **27**, PA4201 (2012).
74. Leese, F., Agrawal, S. & Held, C. Long-distance island hopping without dispersal stages: transportation across major zoogeographic barriers in a Southern Ocean isopod. *Naturwissenschaften* **97**, 583–594 (2010).
75. Fraser, C. I., Nikula, R., Ruzzante, D. E. & Waters, J. M. Poleward bound: biological impacts of Southern Hemisphere glaciation. *Trends Ecol. Evol.* **27**, 462–471 (2012).
76. Thornhill, D. J., Mahon, A. R., Norenburg, J. L. & Halanych, K. M. Open-ocean barriers to dispersal: a test case with the Antarctic Polar Front and the ribbon worm *Parborlasia corrugatus* (Nemertea: Lineidae). *Mol. Ecol.* **17**, 5104–5117 (2008).
77. Poulin, E., González-Wever, C., Díaz, A., Gérard, K. & Hüne, M. Divergence between Antarctic and South American marine invertebrates: what molecular biology



- tells us about Scotia Arc geodynamics and the intensification of the Antarctic Circumpolar Current. *Global Planet. Change* **123**, 392–399 (2014).
78. Page, T. J. & Linse, K. More evidence of speciation and dispersal across the Antarctic Polar Front through molecular systematics of Southern Ocean *Limatula* (Bivalvia: Limidae). *Polar Biol.* **25**, 818–826 (2002).
  79. Wilson, N. G., Schrodl, M. & Halanych, K. M. Ocean barriers and glaciation: evidence for explosive radiation of mitochondrial lineages in the Antarctic sea slug *Doris kerguelensis* (Mollusca, Nudibranchia). *Mol. Ecol.* **18**, 965–984 (2009).
  80. Díaz, A., Féral, J. P., David, B., Saucède, T. & Poulin, E. Evolutionary pathways among shallow and deep-sea echinoids of the genus *Sterechinus* in the Southern Ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **58**, 205–211 (2011).
  81. O'Hara, T. D., Smith, P. J., Mills, V. S., Smirnov, I. & Steinke, D. Biogeographical and phylogeographical relationships of the bathyal ophiuroid fauna of the Macquarie Ridge, Southern Ocean. *Polar Biol.* **36**, 321–333 (2013).
  82. Hunter, R. L. & Halanych, K. M. Evaluating connectivity in the brooding brittle star *Astrota ma agassizii* across the Drake Passage in the Southern Ocean. *J. Hered.* **99**, 137–148 (2008).
  83. Barnes, D. K. A., Hodgson, D. A., Convey, P., Allen, C. S. & Clarke, A. Incursion and excursion of Antarctic biota: past, present and future. *Glob. Ecol. Biogeogr.* **15**, 121–142 (2006).
  84. Huiskes, A. H. L. *et al.* Aliens in Antarctica: assessing transfer of plant propagules by human visitors to reduce invasion risk. *Biol. Conserv.* **171**, 278–284 (2014).
  85. Molina-Montenegro, M. A. *et al.* Assessing the importance of human activities for the establishment of the invasive *Poa annua* in Antarctica. *Polar Res.* **33**, 21425 (2014).
  86. Volonteri, O., de León, R. P., Convey, P. & Krzemińska, E. First record of Trichoceridae (Diptera) in the maritime Antarctic. *Polar Biol.* **36**, 1125–1131 (2013).
  87. Lewis, P. N., Riddle, M. & Hewitt, C. L. Management of exogenous threats to Antarctica and the sub-Antarctic Islands: balancing the risks from TBT and non-indigenous marine organisms. *Mar. Pollut. Bull.* **49**, 999–1005 (2004).  
**This study showed that a diverse fouling community can be transported to Antarctica on the hulls of research vessels, though sea-ice may reduce the numbers of organisms being transported.**
  88. Aronson, R. B., Frederich, M., Price, R. & Thatje, S. Prospects for the return of shell-crushing crabs to Antarctica. *J. Biogeogr.* **42**, 1–7 (2015).
  89. Griffiths, H. J., Whittle, R. J., Roberts, S. J., Belchier, M. & Linse, K. Antarctic crabs: invasion or endurance? *PLoS ONE* **8**, e66981 (2013).
  90. Berkman, P. A., Lang, M. A., Walton, D. W. H. & Young, O. R. (eds) *Science Diplomacy. Antarctica, Science, and the Governance of International Spaces*. (Smithsonian Institution, 2011).
  91. Convention on Biological Diversity. National Biodiversity Strategy and Action Plans. <http://www.cbd.int/nbsap/> (2015).
  92. Shaw, J. D., Terauds, A., Riddle, M. J., Possingham, H. P. & Chown, S. L. Antarctica's protected areas are inadequate, unrepresentative, and at risk. *PLoS Biol.* **12**, e1001888 (2014).
  93. Hughes, K. A. & Convey, P. The protection of Antarctic terrestrial ecosystems from inter- and intra-continental transfer of non-indigenous species by human activities: a review of current systems and practices. *Glob. Environ. Change* **20**, 96–112 (2010).
  94. Braun, C. *et al.* in *Antarctic Futures. Human Engagement with the Antarctic Environment* (eds Tin, T., Liggett, D., Maher, P. T. & Lamers, M.) 169–191 (Springer, 2014).
  95. CEP (Committee for Environmental Protection). Non-native species manual. [http://www.ats.aq/documents/atcm34/ww/atcm34\\_ww004\\_e.pdf](http://www.ats.aq/documents/atcm34/ww/atcm34_ww004_e.pdf) (2011).
  96. Nielsen, U. N. & Wall, D. H. The future of soil invertebrate communities in polar regions: different climate change responses in the Arctic and Antarctic? *Ecol. Lett.* **16**, 409–419 (2013).
  97. Smith, W. O. Jr, Ainley, D. G., Arrigo, K. R. & Dinniman, M. S. The oceanography and ecology of the Ross Sea. *Annu. Rev. Mar. Sci.* **6**, 469–487 (2014).
  98. Ainley, D. G. *et al.* Decadal trends in abundance, size and condition of Antarctic toothfish in McMurdo Sound, Antarctica, 1972–2011. *Fish Fish.* **14**, 343–363 (2013).
  99. Brady, A.-M. (ed.) *The Emerging Politics of Antarctica* (Routledge, 2013).
  100. Puig-Marcó, R. Access and benefit sharing of Antarctica's biological material. *Mar. Genomics* **17**, 73–78 (2014).
  101. Wan, E. *et al.* Green technologies for room temperature nucleic acid storage. *Curr. Issues Mol. Biol.* **12**, 135–142 (2010).
  102. Fretwell, P. T. *et al.* An emperor penguin population estimate: the first global, synoptic survey of a species from space. *PLoS ONE* **7**, e33751 (2012).  
**This study showed that a synoptic survey of the entire population of an important Antarctic species, the Emperor Penguin, can be undertaken for a single year by satellite remote sensing, with the numbers of breeding pairs estimated increasing over previous counts by >50,000.**
  103. Shin, J.-I., Kim, H.-C., Kim, S.-I. & Hong, S. G. Vegetation abundance on the Barton Peninsula, Antarctica: estimation from high-resolution satellite images. *Polar Biol.* **37**, 1579–1588 (2014).
  104. van Dorst, J. *et al.* Community fingerprinting in a sequencing world. *FEMS Microbiol. Ecol.* **89**, 316–330 (2014).
  105. Lynch, H. J., Naveen, R., Trathan, P. N. & Fagan, W. J. Spatially integrated assessment reveals widespread changes in penguin populations on the Antarctic Peninsula. *Ecology* **93**, 1367–1377 (2012).
  106. Peck, L. S. & Clark, M. S. in *Adaptation and Evolution in Marine Environments, Volume 1* (eds di Prisco, G. & Verde, C.) 157–182 (Springer, 2012).
  107. Bednaršek, N. *et al.* Extensive dissolution of live pteropods in the Southern Ocean. *Nat. Geosci.* **5**, 881–885 (2012).
  108. Kawaguchi, S. *et al.* Risk maps for Antarctic krill under projected Southern Ocean acidification. *Nature Climate Change* **3**, 843–847 (2013).
  109. Anson, I. J. & Lutjeharms, J. R. E. Eddies originating at the South-West Indian Ridge. *J. Mar. Syst.* **39**, 1–18 (2003).
  110. Fretwell, P. *et al.* Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. *Cryosphere* **7**, 375–393 (2013).

**Acknowledgements** This work was supported by Australian Research Council grants DP140102815 to S.L.C. and DP150103017 to M.A.M., an Australian Research Council Discovery Early Career Fellowship DE140101715 to C.I.F., grants from the New Zealand Antarctic Research Institute, New Zealand Marsden Fund and the US National Science Foundation to S.C.C., and emeritus support from the British Antarctic Survey to A.C. We thank D. J. Marshall, C. Lee and H. W. Morgan for comments that improved the work.

**Author Contributions** S.L.C., S.C.C. and M.A.M. conceived the work; C.I.F. conceptualized and drew the figures; all authors contributed equally to the planning and writing of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence should be addressed to S.L.C. ([steven.chown@monash.edu](mailto:steven.chown@monash.edu)).

# Competitive binding of antagonistic peptides fine-tunes stomatal patterning

Jin Suk Lee<sup>1,2</sup>, Marketa Hnilova<sup>3</sup>, Michal Maes<sup>2</sup>, Ya-Chen Lisa Lin<sup>1,2</sup>, Aarthi Putarjuna<sup>2</sup>, Soon-Ki Han<sup>1,2</sup>, Julian Avila<sup>1,2</sup> & Keiko U. Torii<sup>1,2</sup>

**During development, cells interpret complex and often conflicting signals to make optimal decisions. Plant stomata, the cellular interface between a plant and the atmosphere, develop according to positional cues, which include a family of secreted peptides called epidermal patterning factors (EPFs). How these signalling peptides orchestrate pattern formation at a molecular level remains unclear. Here we report in *Arabidopsis* that Stomagen (also called EPF-LIKE9) peptide, which promotes stomatal development, requires ERECTA (ER)-family receptor kinases and interferes with the inhibition of stomatal development by the EPIDERMAL PATTERNING FACTOR 2 (EPF2)-ER module. Both EPF2 and Stomagen directly bind to ER and its co-receptor TOO MANY MOUTHS. Stomagen peptide competitively replaced EPF2 binding to ER. Furthermore, application of EPF2, but not Stomagen, elicited rapid phosphorylation of downstream signalling components *in vivo*. Our findings demonstrate how a plant receptor agonist and antagonist define inhibitory and inductive cues to fine-tune tissue patterning on the plant epidermis.**

Development and pattern formation of multicellular organisms rely on diffusible signals that instruct cells to adopt a specific fate for optimal function, and hence organismal fitness. Often such signals are encoded by multiple gene families, which raises the question of how a given cell orchestrates the decision-making process. For instance, a family of secreted signals, such as FGFs, are used in an iterative manner to specify multiple, diverse developmental processes in animals<sup>1</sup>. While peptide signalling has recently emerged as a critical regulator of plant development<sup>2</sup>, how specific members of plant peptide families share and distribute functions remains unclear. Patterning of stomata—valves on the plant epidermis that mediate carbon dioxide acquisition and water control—relies on cell–cell communication, which specifies a subset of seemingly uniform protodermal cells to acquire stomatal progenitor fate. Two secreted cysteine-rich peptides, EPF1 and EPF2, are expressed in later and earlier stages of stomatal precursors, respectively, and are perceived by the cell-surface receptors, ER-family leucine-rich repeat receptor kinases (LRR-RKs)—ER, ER-LIKE1 (ERL1) and ERL2—to inhibit stomatal development<sup>3–7</sup>. The receptor-like protein TOO MANY MOUTHS (TMM) modulates the signalling strengths of ER-family receptor kinases in a region-specific manner<sup>6,8</sup>. Genetic evidence suggests that the signals are mediated via a mitogen-activated protein kinase (MAPK) cascade, which eventually downregulates the transcription factor responsible for initiating stomatal lineage via direct phosphorylation<sup>9–12</sup>.

Recently, EPF-LIKE9 (EPFL9) peptide, also known as Stomagen, was identified as a positive regulator of stomatal development, a role opposite to EPF1 and EPF2 (refs 13–17). Structural modelling of the EPF/EPFL-family peptides using the NMR-solved structure of Stomagen predicts that they all adopt related structures<sup>16</sup>. This raises the question of how can structurally related peptides confer completely opposite developmental responses. The molecular mechanism for Stomagen action remains unknown.

## Stomagen acts downstream of the ER family

To place Stomagen into a genetic framework of the core stomatal signalling pathway, we first examined the effects of induced *STOMAGEN*

overexpression (*iSTOMAGEN*) on *er erl1 erl2* triple mutant phenotypes by an oestradiol-induction system or co-suppression by artificial microRNA (*STOMAGEN-ami*) (Fig. 1 and Extended Data Figs 1–4). As previously reported<sup>13,14</sup>, ectopic *iSTOMAGEN* expression resulted in an increase in stomatal density (number of stomata per mm<sup>2</sup>), stomatal index (percentage of stomata per total number of stomatal and non-stomatal epidermal cells) and severe stomatal clustering in wild-type cotyledon epidermis (Fig. 1a, b, k and Extended Data Figs 1–3). In contrast, *iSTOMAGEN* had no effect on stomatal density, stomatal index, or stomatal clusters in *er erl1 erl2* cotyledons, just like in *tmm* (Fig. 1 and Extended Data Fig. 3)<sup>13,14</sup>, suggesting that *STOMAGEN* and the ER family act in the same pathway.

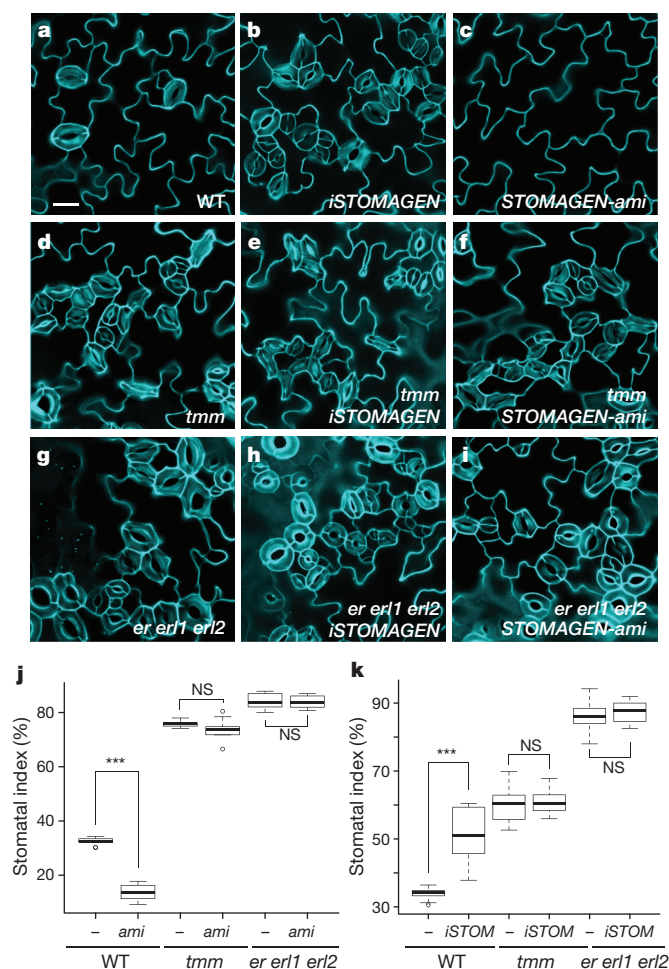
As reported, *STOMAGEN-ami* lines markedly reduced stomatal development in wild-type cotyledons (Fig. 1a, c, j and Extended Data Fig. 4)<sup>13</sup>. In contrast, *STOMAGEN-ami* had no effect on stomatal density, stomatal index, and stomatal clustering phenotype of *er erl1 erl2* cotyledons, just like *tmm* (Fig. 1 and Extended Data Fig. 4). Thus, ER-family receptor kinases are required for the hypermorphic and hypomorphic effects of Stomagen. The epistasis of *er erl1 erl2* stomatal cluster phenotype over the phenotype of *STOMAGEN-ami* places the ER family downstream of *STOMAGEN*, consistent with the molecular identity of their gene products as receptor kinases and a secreted peptide.

## Genetic dissection of Stomagen action

To dissect the role of Stomagen on the TMM and ER module, we comprehensively investigated the effects of *iSTOMAGEN* on stomatal differentiation in *tmm* hypocotyls with additional *er*-family mutations (Fig. 2 and Extended Data Fig. 5). In hypocotyls, TMM and ER family have opposite functions: *tmm* hypocotyls lack stomata<sup>18</sup>, whereas *er erl1 erl2* hypocotyls produce stomatal clusters<sup>19</sup>. While *tmm* is epistatic to *er* single mutation in hypocotyls, consecutive loss of ER-family genes reverts stomatal development in a dosage-dependent manner, with *er erl1 erl2* being epistatic to *tmm* (ref. 7). *iSTOMAGEN* does not confer stomatal differentiation in *tmm*

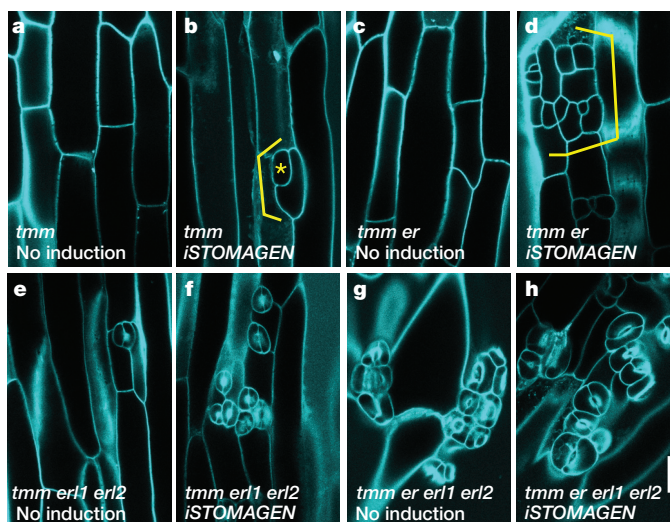
<sup>1</sup>Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, USA. <sup>2</sup>Department of Biology, University of Washington, Seattle, Washington 98195, USA. <sup>3</sup>Department of Materials Science and Engineering, University of Washington, Seattle, Washington 98195, USA.





**Figure 1 | Complete loss of ER-family genes confers insensitivity to STOMAGEN overexpression and co-suppression.** a–i, Representative confocal images of cotyledon abaxial epidermis from 10-day-old light-grown seedlings of wild type (a–c), *tmm* (d–f) and *er erl1 erl2* (g–i), with induced STOMAGEN overexpression (*iSTOMAGEN*; b, e, h) or STOMAGEN-*ami* construct (c, f, i). Uninduced controls show no effects (see Extended Data Figs 1–3). Images were taken under the same magnification. Scale bar, 30  $\mu$ m. *n* = 13 (a); *n* = 18 (b); *n* = 26 (c); *n* = 16 (d); *n* = 24 (e); *n* = 26 (f); *n* = 16 (g); *n* = 24 (h); *n* = 12 (i). j, Stomatal index. –, control; *ami*, STOMAGEN-*ami*. Box, mean; bars, s.e.m. \*\*\**P* < 0.005 (Wilcoxon rank sum test). NS, not significant (*P* = 0.653 for *tmm*; *P* = 0.539 for *er erl1 erl2*). *n* = 8 for each genotype. k, Stomatal index. –, uninduced; *iSTOM*, induced. Box, mean; bars, s.e.m. \*\*\**P* < 0.005 (Wilcoxon rank sum test). NS, not significant (*P* = 0.114 for *tmm*; *P* = 0.688 for *er erl1 erl2*). No induction, *n* = 16; *iSTOM*, *n* = 14; *tmm* no induction, *tmm iSTOM*, *er erl1 erl2*, *er erl1 erl2 iSTOM*, *n* = 15 for each genotype. For the total numbers of stomata counted, see legends for Extended Data Figs 3 and 4.

hypocotyls<sup>13</sup>. However, in some instances arrested stomatal precursor cells (stomatal-lineage ground cells (SLGCs)) were observed, indicating that, in the absence of TMM, *iSTOMAGEN* could initiate stomatal development in hypocotyls (Fig. 2a, b and Extended Data Fig. 5c, d). Additional *er*-family mutations exaggerated this effect; that is, *iSTOMAGEN* in *tmm er* and *tmm erl1 erl2* hypocotyls, both of which lack stomata, resulted in SLGC clusters (Fig. 2c, d and Extended Data Fig. 5e–h). *iSTOMAGEN* triggered stomatal cluster formation in *tmm erl1*, *tmm erl1 erl2* and *tmm er erl1* mutant hypocotyls, while intensifying stomatal entry divisions in *tmm er erl2* hypocotyls (Fig. 2e, f and Extended Data Fig. 5i–p). Different effects of *iSTOMAGEN* on the higher-order mutants lacking ER (for example, *tmm er* and *tmm er erl2*) compared with those lacking ERL1 (for example, *tmm erl1* and *tmm erl1 erl2*) reflect the overlapping yet unique roles of ER and ERL1

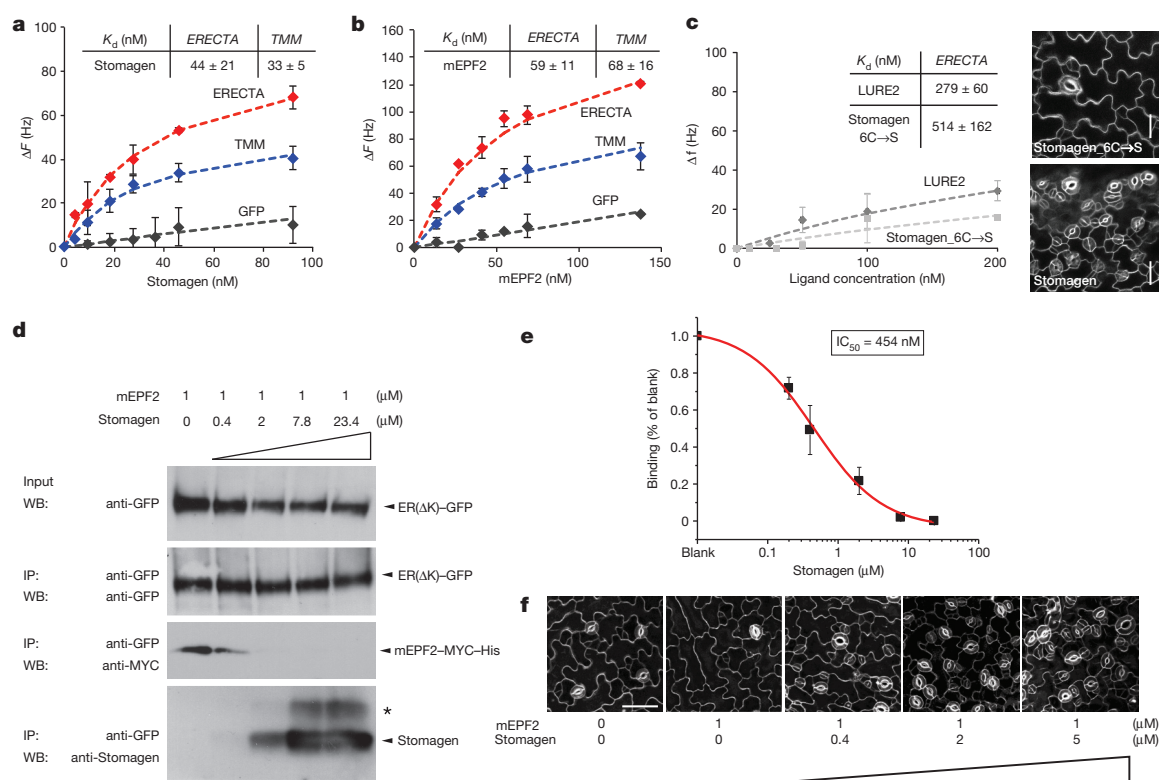


**Figure 2 | STOMAGEN overexpression on stomatal development in *tmm* hypocotyl epidermis with combinatorial loss-of-function in ER-family genes.** a–h, Representative confocal microscopy images of hypocotyl epidermis from 10-day-old light-grown transgenic *Est::STOMAGEN* (oestradiol-induced STOMAGEN) seedlings of *tmm* (a, b); *tmm er* (c, d); *tmm erl1 erl2* (e, f); and *tmm er erl1 erl2* (g, h). A control, uninduced phenotype (a, c, e, g) and *iSTOMAGEN* phenotype (b, d, f, h) is shown. *iSTOMAGEN* results in arrested stomatal precursor cells (asterisk) and stomatal-lineage ground cells (SLGCs (bracket)) in *tmm* hypocotyls (b). Additional *er* mutation exaggerated this effect (d), while additional *erl1 erl2* mutations increased stomata (f). Images were taken under the same magnification. Scale bar, 30  $\mu$ m. *n* = 20 (a); *n* = 20 (b); *n* = 19 (c); *n* = 22 (d); *n* = 17 (e); *n* = 20 (f); *n* = 20 (g); *n* = 20 (h). For a complete set of higher-order mutant phenotypes and quantitative data, see Extended Data Fig. 5.

in stomatal development<sup>6</sup>. Finally, *iSTOMAGEN* failed to enhance the severe stomatal clustering phenotype in *tmm er erl1 erl2* mutants (Fig. 2g, h and Extended Data Fig. 5q, r). Quantitative analysis of stomatal index and SLGC index (the percentage of SLGCs in total epidermal cells) supports these findings (Extended Data Fig. 5s, t). Together, the results suggest that in the hypocotyls, where TMM and ER-family act antagonistically, Stomagen primarily acts via all three ER-family receptor kinases.

Among the ER family, ER primarily perceives EPF2 to restrict initiation of stomatal cell lineages, while ERL1 primarily perceives EPF1 to orient stomatal spacing and prevent guard cell differentiation<sup>6</sup>. As such, *epf2* increases SLGCs, whereas *epf1* violates stomatal spacing<sup>3–5</sup>. Neither *epf2* nor *epf1* confers severe stomatal clustering phenotype like *iSTOMAGEN*, since only a subset of ER-family-mediated pathways has been compromised<sup>6</sup>. We delineated the role of Stomagen in each of these steps. We first examined whether EPF1, EPF2 and STOMAGEN transcripts are under feedback regulation, which may complicate the genetic analyses. EPF1 and EPF2 transcript levels were slightly upregulated by *iSTOMAGEN*, and conversely, slightly downregulated by STOMAGEN-*ami* (Extended Data Fig. 2c, d). However, the endogenous STOMAGEN transcript levels are unaffected by *epf1*, *epf2*, or *epf1 epf2* (Extended Data Fig. 2d). Thus, altered expression of EPF1 and EPF2 by STOMAGEN misregulation probably reflects the numbers of stomatal-lineage cells<sup>13,14</sup>.

EPF2–ER or EPF1–ERL1 signalling pathways with *iSTOMAGEN* resulted in severe stomatal clusters, indicating that excessive Stomagen promotes stomatal differentiation when either pathway is compromised (Extended Data Fig. 3). These genetic data support the notion that Stomagen, when ectopically overexpressed, can bind to all ER-family receptor kinases and inhibit signal transduction. Indeed, co-immunoprecipitation experiments using *Nicotiana benthamiana* microsomal fraction expressing green fluorescent protein (GFP)-fused ectodomains of ER, ERL1, ERL2 or TMM incubated



**Figure 3 | Direct and competitive binding of Stomagen and EPF2 peptides to ER.** **a–c**, QCM analysis for direct binding. **a**, **b**, The averages of experimental frequency shift values recorded from two to four independent experiments for Stomagen (**a**) or mEPF2 (**b**) onto biosensor chips functionalized with ER( $\Delta$ K)-GFP (red), TMM-GFP (blue) and GFP alone (grey) and fitted to the Langmuir adsorption model using least square regression. Error bars indicate s.d. Stomagen-ER,  $n = 3$ ; Stomagen-TMM,  $n = 2$ ; Stomagen-GFP,  $n = 3$ ; mEPF2-ER,  $n = 2$ ; mEPF2-TMM,  $n = 3$ ; mEPF2-GFP,  $n = 2$ . **c**, The average experimental frequency shift values recorded for LURE2 (dark grey) and mutant Stomagen (light grey) on ER( $\Delta$ K)-GFP. To calculate the dissociation constant ( $K_d$ ) values, the ligand concentrations were increased to 1  $\mu$ M to obtain fitted curves. See Extended Data Fig. 9 for raw recording data. Inactive mutant Stomagen, in which six cysteines were substituted with serines (Stomagen\_C6→S)-ER,  $n = 3$ ; LURE2-ER,  $n = 2$ . Error bars indicate s.d. Right insets: wild-type cotyledon epidermis treated with 2.5  $\mu$ M mutant or

bioactive Stomagen. Scale bars, 30  $\mu$ m.  $n = 8$  for each treatment. For **a–c**, each experimental point in independent binding experiment (referred to as ' $n = 1$ ') is generated from multi-point raw data (10–20 points) with average and s.d. values. **d**, Competitive binding. Microsomal fractions expressing ER( $\Delta$ K)-GFP were incubated with 1  $\mu$ M of bioactive mEPF2 with increasing concentrations of bioactive Stomagen and subjected to immunoprecipitation. The mEPF2-MYC-His blot was re-probed with anti-Stomagen antibody. Asterisk indicates most likely isomer. **e**, Quantitative analysis of competition from four biological replicates. Error bars, s.e.m. The IC<sub>50</sub> value is substantially higher than the  $K_d$  values for Stomagen-ER or mEPF2-ER, presumably owing to the immunoblot-based quantification. **f**, Wild-type cotyledon epidermis treated with mEPF2 alone or simultaneously co-treated with mEPF2 and increasing concentrations of Stomagen for 5 days.  $n = 3$  for each treatment. Images were taken under the same magnification. Scale bar, 50  $\mu$ m.

with synthetic Stomagen peptides demonstrated that Stomagen associates with all ER-family receptor kinases and TMM (Extended Data Fig. 6a).

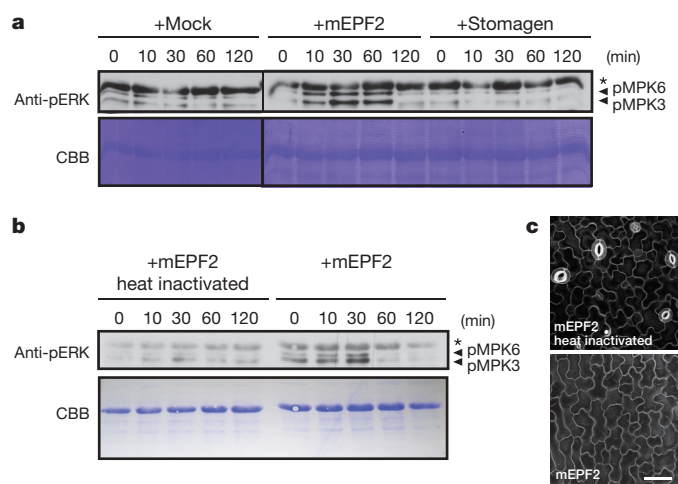
Unlike overexpression, Stomagen co-suppression imposed different effects on EPF2-ER and EPF1-ERL1 signalling pathways. *STOMAGEN-ami* suppressed the stomatal-pairing phenotype of *epf1* and dominant-negative, kinases-deleted ERL1 (ERL1( $\Delta$ K)) *erl1* (Extended Data Fig. 4g–j, m). In contrast, *STOMAGEN-ami* exhibited complex interactions with *epf2* and dominant-negative ER (ER( $\Delta$ K)) *er*, reducing numbers of stomata but not that of SLGCs (Extended Data Fig. 4c–f, k–n). This supports the idea that Stomagen counteracts EPF2 for ER-mediated stomatal initiation<sup>13,14,16</sup>. This also suggests that, in the absence of both *EPF2* and *STOMAGEN*, the default ER pathway is not activated while the later ERL-mediated pathway remains capable of repressing the differentiation of mature stomata.

### Competitive binding of EPF2 and Stomagen

A series of genetic analyses leads to the possibility that Stomagen antagonizes EPF2's action via direct binding to ER. To address this, we produced bioactive Stomagen and predicted mature EPF2 (mEPF2) peptides (Extended Data Figs 7 and 8). Subsequently, their direct binding to ER as well as to TMM was tested using previously

established quartz crystal microbalance (QCM) biosensor platforms (Fig. 3a, b and Extended Data Fig. 9)<sup>6</sup>. Briefly, we immobilized purified GFP-fused receptors or control GFP from *N. benthamiana* on gold surfaces of QCM chips via anti-GFP antibody and then introduced the bioactive Stomagen or mEPF2 peptide solutions. The peptide-receptor binding was recorded as a function of frequency change (see Methods)<sup>6</sup>. Both Stomagen and mEPF2 exhibited saturable binding to the ER ectodomain fused to GFP (ER( $\Delta$ K)-GFP) with similar dissociation constants at a nanomolar range (Fig. 3a, b and Extended Data Fig. 9). Additionally, Stomagen and mEPF2 bound to TMM with high affinity (Fig. 3a, b). No significant binding of Stomagen or mEPF2 to control GFP was detected (Fig. 3a, b and Extended Data Fig. 9). To address the specificity of peptide-receptor interactions, two control peptides were subjected to the QCM analysis using ER( $\Delta$ K)-GFP-functionalized chips: (1) non-folding, inactive mutant Stomagen, in which six cysteines were substituted with serines (Extended Data Fig. 8g)<sup>16</sup>; and (2) LURE2, an unrelated cysteine-rich peptide, which acts as a pollen-tube attractant<sup>20</sup>. Neither mutant Stomagen nor LURE2 exhibited binding above the background level (Fig. 3c). Consistently, LURE2 did not associate with ER, TMM, or an innate immunity receptor FLS2 (refs 21, 22) fused with GFP expressed in *N. benthamiana* in co-immunoprecipitation assays (Extended Data





**Figure 4 | EPF2, but not Stomagen, triggers downstream MAPK activation in *Arabidopsis* seedlings.** **a**, **b**, Differential MAPK activation in *Arabidopsis* wild-type seedlings treated with buffer only (**a**, mock), mEPF2 (**a**, **b**), Stomagen (**a**), and heat-denatured mEPF2 (**b**) for respective time intervals (min). The blots were probed with anti-phospho-ERK antibody (anti-pERK) to detect phosphorylated MPK6 (pMPK6) and pMPK3 upon peptide treatment. Asterisk indicates non-specific band. CBB, total proteins stained. Four and two biological replicates were performed for **a** and **b**, respectively. **c**, Confocal microscopy of *Arabidopsis* wild-type cotyledon abaxial epidermis treated with heat-denatured mEPF2 (top) and control, non-denatured mEPF2 (bottom). Scale bar, 40  $\mu$ m.  $n = 3$  for each treatment.

Fig. 6b). Similarly, FLS2 failed to immunoprecipitate Stomagen above background levels (Extended Data Fig. 6c). Together, the results emphasize the specificity of Stomagen-ERECTA and Stomagen-TMM interactions.

Next, we performed ligand competition assays between Stomagen and EPF2. Microsomal fractions from *N. benthamiana* expressing ER ectodomain (ER( $\Delta$ K)-GFP) were incubated with bioactive epitope-tagged mEPF2 (mEPF2-MYC-HIS; 1  $\mu$ M) and increasing concentrations of bioactive Stomagen peptides (0–23.4  $\mu$ M) followed by immunoprecipitation of ER. Co-immunoprecipitated epitope-tagged mEPF2 was detected first. Then, the same blot was re-probed with anti-Stomagen antibody to detect co-immunoprecipitated Stomagen. Increasing concentrations of Stomagen peptide replaced mEPF2 for ER binding (Fig. 3d). Quantitative analysis confirmed the competitive binding of Stomagen and mEPF2 to ER, with a half-maximum inhibitory concentration ( $IC_{50}$ ) value of 454 nM (Fig. 3e). Our results demonstrate that Stomagen and EPF2 peptides directly compete for binding to the same receptor, ER. Application of mEPF2 to wild-type seedlings inhibited stomatal development, while simultaneous treatment of mEPF2 with increasing concentration of Stomagen in a similar concentration range used in the competition experiments resulted in increased stomatal differentiation (Fig. 3f). The results are consistent with a previous report<sup>16</sup> and further emphasize the *in vivo* biological relevance of peptide competition.

### Activation of downstream signalling

To unravel the mechanism of Stomagen as a competitive antagonist of EPF2, we examined the activation of downstream signalling, specifically, using phosphorylation of MPK3 and MPK6 as readout. Genetic studies suggest that EPF2-ER ligand-receptor signalling acts via a MAPK cascade<sup>9–12</sup>. However, a recent report of co-expressed stomatal signalling components in *N. benthamiana* failed to detect MPK6 activation by EPF2 (ref. 17), probably due to a limitation of the heterologous co-expression system for capturing fast and transient response. We therefore tested MAPK activation *in vivo* using *Arabidopsis* seedlings. Application of mEPF2 peptide to *Arabidopsis* wild-type seedlings rapidly elicited phosphorylation of MPK3 and

MPK6 in 10 min, a characteristic signature of MAPK activation, which declined after 2 h (Fig. 4a, b). Heat-induced denaturation of mEPF2 greatly diminished MAPK phosphorylation, correlating with its loss of bioactivity (Fig. 4b, c). By contrast, Stomagen peptide treatment failed to trigger MAPK phosphorylation (Fig. 4a). We conclude that EPF2 activates ER signalling, leading to subsequent MAPK activation to inhibit stomatal development, while Stomagen prevents the signal transduction.

### Discussion

Our work elucidates the competitive binding of Stomagen and EPF2 to ER as a molecular mechanism optimizing stomatal patterning. Plant genomes possess large numbers of peptide gene families, the functions of which remain largely unknown<sup>23</sup>. The concept of fine-tuning signal transduction by related endogenous peptides that assume opposing functions may extend to other peptide families. EPF2 is expressed in a subset of protodermal cells, while Stomagen is secreted from an underlying internal tissue<sup>4,5,13,14</sup>. Thus, it seems plausible that a protodermal cell might respond to differences in intrinsic concentrations of EPF2 and Stomagen on each neighbouring side. It remains to be tested whether local concentrations of Stomagen in the apoplast reflect the  $IC_{50}$  values we have determined biochemically (Fig. 3e). The complex effects of *STOMAGEN* overexpression on a series of *er*-family mutants in the *tmm* background (Fig. 2) resemble that of *challah* (*chal*) higher-order mutants, which lack EPFL4 and EPFL6 peptides, another set of ER ligands promoting stem growth<sup>24–26</sup>. This raises the possibility that complex fine-tuning of multiple EPF-family peptides may occur at developmental contexts far beyond stomatal patterning. Quantitative visualization of each peptide *in vivo* during epidermal development, as well as precise documentation of the dose-response effects of simultaneous mixed peptide applications of wide concentration gradients, may reveal the signalling complexity at the level of ligand-receptor association. EPF2 and Stomagen bind to ER and TMM with a similar affinity (Fig. 3), suggesting the formation of co-receptor complexes, a hallmark of receptor activation and signal transduction in plant LRR-RKs in development and innate immunity response<sup>27,28</sup>. Future structural and cell biological studies may reveal the intricacy behind how a cell interprets conflicting signals to make decisions during developmental patterning.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 5 August 2014; accepted 14 May 2015.

Published online 17 June 2015.

- Dorey, K. & Amaya, E. FGF signalling: diverse roles during early vertebrate embryogenesis. *Development* **137**, 3731–3742 (2010).
- Fukuda, H. & Higashiyama, T. Diverse functions of plant peptides: entering a new phase. *Plant Cell Physiol.* **52**, 1–4 (2011).
- Hara, K., Kajita, R., Torii, K. U., Bergmann, D. C. & Kakimoto, T. The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. *Genes Dev.* **21**, 1720–1725 (2007).
- Hara, K. *et al.* Epidermal cell density is auto-regulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in *Arabidopsis* leaves. *Plant Cell Physiol.* **50**, 1019–1031 (2009).
- Hunt, L. & Gray, J. E. The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Curr. Biol.* **19**, 864–869 (2009).
- Lee, J. S. *et al.* Direct interaction of ligand-receptor pairs specifying stomatal patterning. *Genes Dev.* **26**, 126–136 (2012).
- Shpak, E. D., McAbee, J. M., Pillitteri, L. J. & Torii, K. U. Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science* **309**, 290–293 (2005).
- Nadeau, J. A. & Sack, F. D. Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science* **296**, 1697–1700 (2002).
- Wang, H., Ngwenyama, N., Liu, Y., Walker, J. & Zhang, S. Stomatal development and patterning are regulated by environmentally responsive mitogen-activated protein kinases in *Arabidopsis*. *Plant Cell* **19**, 63–73 (2007).
- Lampard, G. R., Macalister, C. A. & Bergmann, D. C. *Arabidopsis* stomatal initiation is controlled by MAPK-mediated regulation of the bHLH SPEECHLESS. *Science* **322**, 1113–1116 (2008).

11. Bemis, S. M., Lee, J. S., Shpak, E. D. & Torii, K. U. Regulation of floral patterning and organ identity by *Arabidopsis* ERECTA-family receptor kinase genes. *J. Exp. Bot.* **64**, 5323–5333 (2013).
12. Bergmann, D. C., Lukowitz, W. & Somerville, C. R. Stomatal development and pattern controlled by a MAPKK kinase. *Science* **304**, 1494–1497 (2004).
13. Sugano, S. S. *et al.* Stomagen positively regulates stomatal density in *Arabidopsis*. *Nature* **463**, 241–244 (2010).
14. Kondo, T. *et al.* Stomatal density is controlled by a mesophyll-derived signaling molecule. *Plant Cell Physiol.* **51**, 1–8 (2010).
15. Hunt, L., Bailey, K. J. & Gray, J. E. The signalling peptide EPFL9 is a positive regulator of stomatal development. *New Phytol.* **186**, 609–614 (2010).
16. Ohki, S., Takeuchi, M. & Mori, M. The NMR structure of stomagen reveals the basis of stomatal density regulation by plant peptide hormones. *Nat. Commun.* **2**, 512 (2011).
17. Jewaria, P. K. *et al.* Differential effects of the peptides Stomagen, EPF1 and EPF2 on activation of MAP kinase MPK6 and the SPCH protein level. *Plant Cell Physiol.* **54**, 1253–1262 (2013).
18. Geisler, M., Yang, M. & Sack, F. D. Divergent regulation of stomatal initiation and patterning in organ and suborgan regions of the *Arabidopsis* mutants *too many mouths* and *four lips*. *Planta* **205**, 522–530 (1998).
19. Pillitteri, L. J., Bogenschutz, N. L. & Torii, K. U. The bHLH protein, MUTE, controls differentiation of stomata and the hydathode pore in *Arabidopsis*. *Plant Cell Physiol.* **49**, 934–943 (2008).
20. Okuda, S. *et al.* Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. *Nature* **458**, 357–361 (2009).
21. Gómez-Gómez, L. & Boller, T. FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Mol. Cell* **5**, 1003–1011 (2000).
22. Zipfel, C. *et al.* Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* **428**, 764–767 (2004).
23. Hanada, K. *et al.* Small open reading frames associated with morphogenesis are hidden in plant genomes. *Proc. Natl Acad. Sci. USA* **110**, 2395–2400 (2013).
24. Abrash, E. B. & Bergmann, D. C. Regional specification of stomatal production by the putative ligand CHALLAH. *Development* **137**, 447–455 (2010).
25. Abrash, E. B., Davies, K. A. & Bergmann, D. C. Generation of signaling specificity in *Arabidopsis* by spatially restricted buffering of ligand-receptor interactions. *Plant Cell* **23**, 2864–2879 (2011).
26. Uchida, N. *et al.* Regulation of inflorescence architecture by intertissue layer ligand-receptor communication between. *Proc. Natl Acad. Sci. USA* **109**, 6337–6342 (2012).
27. Santiago, J., Henzler, C. & Hothorn, M. Molecular mechanism for plant steroid receptor activation by somatic embryogenesis co-receptor kinases. *Science* **341**, 889–892 (2013).
28. Sun, Y. *et al.* Structural basis for flg22-induced activation of the *Arabidopsis* FLS2–BAK1 immune complex. *Science* **342**, 624–628 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank I. Hara-Nishimura for *STOMAGEN-ami* lines and anti-Stomagen antibody; K. Peterson for *iSTOMAGEN* construct and transgenic lines; M. Kanaoka and N. Kamiya for LURE2 peptides; D. Baulcombe for p19 plasmid; C. Tamerler and M. Sarikaya for use of the HPLC, QCM and MALDI-ToF equipment; A. Hofstetter for technical assistance; and J. McAbee, K. Peterson, T. Imaizumi, B. Wakimoto, S. Di Rubbo and R. Horst for comments. K.U.T. is an HHMI-GBMF Investigator and an Endowed Distinguished Professor of Biology; J.S.L. was an NSERC Postdoctoral Fellow. Y.-C.L.L. was a Mary Gates Undergraduate Research Fellow of the University of Washington.

**Author Contributions** J.S.L. and K.U.T. conceived the project. J.S.L., M.H., M.M., J.A. and Y.-C.L.L. purified peptides and performed ligand–receptor binding and bioassays. J.S.L. and S.-K.H. performed RT–PCR. J.S.L. and A.P. performed MAPK assays. J.S.L. and Y.-C.L.L. performed quantitative analysis of stomatal phenotypes. K.U.T. constructed *STOMAGEN* cDNA plasmid. K.U.T., J.S.L., M.H., M.M., Y.-C.L.L., A.P. and S.-K.H. analysed the data. K.U.T. wrote the manuscript with inputs from all co-authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.U.T. ([ktorii@u.washington.edu](mailto:ktorii@u.washington.edu)).



## METHODS

No statistical methods were used to predetermine sample size.

**Plant materials and growth conditions.** The *Arabidopsis* accession Columbia (Col) was used as wild type. All plants used in this study are in the Col background. The following mutants and transgenic plant lines were reported previously: *er-105*, *erl1-2*, *erl2-1* and their higher order mutants<sup>7</sup>; *tmm-1* (ref. 8); *epf1*, *epf2*, *epf1 epf2* and *tmm-KO* (refs 3, 4); *ER(ΔK)* in *er-105*, *ERL1(ΔK)* in *erl1-2* (ref. 6); and *STOMAGEN-ami* line (line 12; a gift from I. Hara-Nishimura)<sup>13</sup>. The *STOMAGEN-ami* lines were introduced into various mutant or transgenic backgrounds via genetic crosses. Plants were grown as described previously<sup>29</sup>. For phenotypic analysis, seedlings were grown on 0.5 Murashige and Skoog (MS) media containing 1× Gamborg Vitamin (Sigma), 0.75% Bacto Agar, and 1% w/v sucrose under continuous light condition at 50 μmol s<sup>-1</sup> m<sup>-2</sup> (light intensity measured by LI-250A; LI-COR).

**Plasmid construction and transgenic plants generation.** pKUT608 (pENTR-*STOMAGEN*) and pKMP127 (*Est::STOMAGEN*) were generated. See Extended Data Table 1 for plasmid and primer sequence information. Transgenic *Arabidopsis* plants were generated by the floral dip method<sup>30</sup>. Multiple transgenic lines per construct were subjected to phenotypic characterization and representative lines (three lines if lines were established, and 12–14 lines if T1 lines were used) were used for quantitative analyses. The *Est::STOMAGEN* lines were introduced into various mutants or transgenic backgrounds via genetic crosses.

**Chemical induction of transgene.** Transgenic *Arabidopsis* seedlings carrying *Est::STOMAGEN* was germinated on 0.5 MS medium supplemented with 10 μM oestradiol (Sigma). Induction of *STOMAGEN* gene expression (*iSTOMAGEN*) was confirmed by RT-PCR (see Extended Data Fig. 2). The induction was further confirmed by observing the epidermal phenotypes of cotyledons and hypocotyls using a confocal microscope.

**RT-PCR analysis.** RNA extraction, cDNA synthesis and RT-PCR were performed as previously described<sup>31</sup>. For a list of primers, see Extended Data Table 1.

**Histology, microscopy and image analysis.** Confocal microscopy images were taken using either Zeiss LSM700 operated by Zen2009 (Zeiss) described previously<sup>6</sup> or Leica SP5-WLL operated by LAS AF (Leica). Cell outlines were visualized with either propidium iodide (PI; Molecular Probes) or FM4-64 (Invitrogen) and observed using the HyD detector with excitation 515 nm, emission 623–642 nm. The images were false coloured using Photoshop CS6 (Adobe). Clearing of seedlings by chloral hydrate and observation using differential interference contrast (DIC) microscope was performed as described previously<sup>32</sup>. For histological analysis, seedlings were stained with toluidine blue-O (TBO; Sigma) as follows. Briefly, samples were placed in 9:1 v/v ethanol to acetic acid overnight, rehydrated through reduced ethanol series to deionized water, then stained with 0.5% TBO for 3 min. Seedlings were immediately rinsed with deionized water and subsequently mounted in 15% v/v glycerol. For bright-field and DIC microscopy, images were taken under Olympus BX51 equipped with DP73 digital camera operated by CellSens Standard software (Olympus).

**Quantitative analysis of epidermis.** Abaxial cotyledons from 10-day-old seedlings of respected genotypes were subjected to TBO staining or DIC microscopy. The central regions overlying the distal vascular loop were imaged and numbers of epidermal cells, stomata and their cluster size were quantified. For each genotype, sample size of 14–16 was used and over 1,000 epidermal cells were counted to provide statistical robustness. For cotyledons of *Est::STOMAGEN* lines, individual T1 seedlings were subjected to analysis. For hypocotyls, three representative T2 *Est::STOMAGEN* lines were analysed. For each seedling, a representative image was taken at the exact location to minimize the variance. Specific numbers of stomata are listed for each genotype in corresponding figure legends. Statistical analysis (Wilcoxon rank sum test) was performed using R version 3.0.3 for stomatal density, stomatal index and SLGC index. *P* values are indicated in each figure legend.

**Transient protein expression in *Nicotiana benthamiana*.** *Agrobacterium tumefaciens* strain GV3101 was transformed with expression clones and grown in yeast extract and beef medium supplemented with relevant antibiotics. Bacterial cultures were precipitated and resuspended in infiltration medium (10 mM MgCl<sub>2</sub>, 10 mM MES (pH 5.6) and 150 μM acetosyringone). Culture densities were adjusted to an OD<sub>600</sub> of 1.0, and the cells were incubated at room temperature for 4 h before infiltration. Equal volumes of cultures carrying different constructs were mixed. To enhance transient expression in tobacco, the silencing suppressor p19 (a gift from D. Baulcombe) was co-infiltrated<sup>33</sup>. The bacterial suspensions were infiltrated into young but fully expanded leaves of *N. benthamiana* plants. After infiltration, plants were cultivated at 25 °C and collected for further biochemical assays after 48–72 h.

**Peptide expression, purification and refolding.** Recombinant mEPF2 peptide was prepared as reported previously<sup>6</sup> and the mature Stomagen peptide, either

wild-type or non-refolding mutant version in which all six cysteines were substituted by serines, were chemically synthesized (Invitrogen and BioSynthesis). The Stomagen peptide was dissolved in 20 mM Tris-HCl, pH 8.8, and 50 mM NaCl and refolded (Mini dialysis kit, MWCO:1,000, GE Healthcare) for 3 d at 4 °C using glutathione (reduced and oxidized forms; Wako) and L-arginine ethyl ester dihydrochloride (Sigma). The peptides were further dialysed three times against 50 mM Tris-HCl, pH 8.0 for 1.5 d to remove glutathione. For non-folding mutant Stomagen, chemically synthesized peptides were dissolved in 50 mM Tris-HCl, pH 8.0. The resulting mEPF2, Stomagen and mutant Stomagen peptides were fractionated using C18 column (Gemini) by HPLC (Waters Delta Prep 3000 HPLC) as previously described to determine the purity of each peptide<sup>6</sup>. The separated peaks were collected, and each peak was identified by MALDI-TOF mass spectrometry on an Autoflex II mass spectrometer in positive ion mode (Bruker Daltonics) using 2:1 *α*-cyano-4-hydroxycinnamic acid and 2,5-dihydroxy-benzoic acid matrix. The collected HPLC-purified mEPF2 and Stomagen peptide peaks were freeze-dried, then re-dissolved to appropriate concentration. Quantification of the active populations of peptides was determined using NanoDrop8000 (Thermo Scientific) using the following molar extinction coefficients: Stomagen, 5,960; EPF2, 6,460; LURE2, 23,950 mol<sup>-1</sup> cm<sup>-1</sup>. For bioassays, freeze-dried peptides were re-dissolved to appropriate concentration in MS medium. For subsequent biochemical assays, the amounts of bioactive peptides were calculated from this quantification.

**Peptide bioassays.** Refolded recombinant mEPF2 and Stomagen peptides in buffer were applied to 1-day-old *Arabidopsis* plants that had germinated on 0.5 MS medium. After 5 d of further incubation in 0.5 MS liquid medium containing each peptide (2.5–5 μM concentration), stomatal phenotypes of abaxial cotyledon epidermis were determined by inspection with a confocal microscope as described previously<sup>6</sup>.

**Immunoprecipitation, protein gel electrophoresis and immunoblots.** For immunoprecipitation and co-immunoprecipitation assays, *N. benthamiana* leaves expressing CaMV35S::ERΔK-GFP, CaMV35S::FLS2(ΔK)-GFP, CaMV35S::TMM-GFP, CaMV35S::GFP, or empty vector were subjected to protein preparation (microsomal fraction enrichment for all except soluble GFP). Co-incubation with Stomagen (1 μM) or LURE2 peptides (1 μM) and immunoprecipitation procedure are described in the Ligand Competition Assays section below. Immunoprecipitation using either anti-GFP (Abcam ab290) antibodies and protein gel immunoblot (western blot) analysis using anti-GFP (Invitrogen C163), anti-Flag (Sigma-Aldrich M2), anti-His (Qiagen anti-His5 34660), and anti-Stomagen (a gift from I. Hara-Nishimura)<sup>13</sup> antibodies were performed as described previously<sup>6</sup>. As secondary antibodies, either goat anti-mouse (GE Healthcare NA931) or anti-rabbit IgG horseradish peroxidase-linked antibodies (Sigma A6154) were used at a dilution of 1:50,000. The protein blots were visualized using SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific).

**Quartz crystal microbalance.** QCM measurements were performed using QCM-Z500 (KSV Instruments) and commercially available AT-cut polished QCM crystals with a fundamental resonant frequency of 4.95 MHz (International Crystal Manufacturing Co.) as reported previously<sup>6</sup>. The QCM crystal chips were pre-treated with anti-GFP antibody (Abcam ab290) to functionalize the chip surface. Subsequently, GFP-tagged receptors or GFP expressed in *N. benthamiana* and extracted as a microsomal fraction (for ER(ΔK)-GFP and TMM-GFP) or a total fraction (GFP) were immobilized onto a QCM sensor chip via anti-GFP antibody linkage. The chips were washed with a phosphate buffer extensively. After establishing a stable baseline using phosphate buffer solution, purified bioactive mEPF2 or Stomagen peptides in the phosphate buffer was added stepwise to the QCM chamber. The frequency change for QCM was monitored until no further change was observed, indicating equilibrium. All experiments were performed at 4 °C in stop-flow mode. The peptide–receptor binding was quantified via QCM by measuring the frequency shifts, Δ*F*, at several peptide concentrations. To determine the dissociation constant (*K<sub>d</sub>*) of each peptide–receptor pair, the experimental frequency shift values were fitted to the Langmuir adsorption model:  $-\Delta F = \Delta F_{\max} C / (C + K_d)$ , where Δ*F*<sub>max</sub> is the frequency shift when the binding is saturated and *C* is the concentration of the bulk solution, using a least squares regression.

**Ligand competition assays.** *N. benthamiana* leaves expressing CaMV35S::ER(ΔK)-GFP were ground in liquid nitrogen and homogenized in extraction buffer (100 mM Tris-HCl pH 8.8, 150 mM NaCl, 1 mM EDTA, 20% glycerol, 20 mM NaF, 1 mM PMSF, 1:1,000 Complete protease inhibitor cocktail (Roche Applied Science)). The slurry was centrifuged at 10,000g for 15 min at 4 °C. The supernatant was sonicated on ice and then centrifuged at 100,000g for 30 min at 4 °C to give a pellet of the microsomal fractions. The pellet was resuspended in membrane solubilization buffer (100 mM Tris-HCl at pH 7.3, 150 mM NaCl, 1 mM EDTA, 10% glycerol, 1% Triton X-100, 20 mM NaF, 1 mM

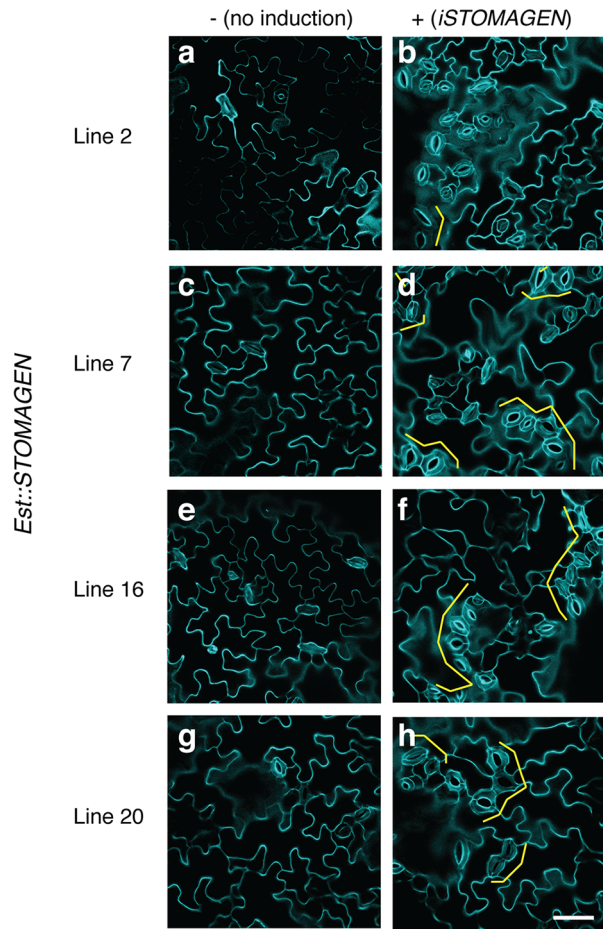
PMSF, 1:1,000 Complete protease inhibitor cocktail) to release membrane proteins. The solution was sonicated on ice and centrifuged again at 100,000g for 30 min at 4 °C. The supernatant was incubated with Protein-G-coupled magnetic beads (Dynabeads Protein G, Invitrogen) that captured anti-GFP (ab290; Abcam) antibody at 4 °C for 2 h with gentle agitation. Then, the beads were washed four times with 500 µl of phosphate buffer (pH 7.4). The immunoprecipitates were suspended in 500 µl of binding buffer (50 mM MES-KOH, pH 5.5 with 100 mM sucrose) containing 1 µM mEPF2-MYC-HIS peptide in the absence or presence of different concentration of unlabelled bioactive Stomagen peptide and then incubated at 4 °C for 1 h with gentle agitation. The reaction mixture was washed four times with 500 µl of phosphate buffer (pH 7.4) to separate bound and free mEPF2-MYC-HIS peptide, and precipitated proteins were eluted with 2× SDS sample buffer at 80 °C for 5 min. Either total membrane or immunoprecipitated proteins were separated on a SDS-PAGE gel and transferred to PDVF membrane (Millipore) for immunoblot analysis using monoclonal anti-GFP (C163, 1:1,000, Invitrogen), anti-MYC (ab32, 1:1,000, abcam) or anti-Stomagen antibodies (1: 5,000, a gift from I. Hara-Nishimura)<sup>13</sup> as primary antibodies. As secondary antibodies, either goat anti-mouse or rabbit IgG horseradish peroxidase-linked antibodies (GE Healthcare NA931; Sigma A6154) were used at a dilution of 1:50,000. Co-immunoprecipitated mEPF2 was detected first. Then, the same blot was re-probed with anti-Stomagen antibody to detect Stomagen.

Four biological replicates were performed and subjected to quantification of the IC<sub>50</sub> values as the following. Band intensities on western blots were quantified using IMAGEJ (<http://rsb.info.nih.gov/ij/index.html>). Pixel values were measured on equal-sized areas and normalized against the bands detecting same immunoprecipitates by monoclonal anti-GFP antibody. The intensity values shown in the paper are the ratios relative to the references, and values were analysed by nonlinear regression analysis using OriginLab version 6 (OriginLab) to calculate the IC<sub>50</sub> value.

**MAPK phosphorylation assays.** 12-day-old *Arabidopsis* seedlings were grown for 5 days on 0.5 MS media plates and then transferred to 0.5 MS liquid media in a 12-well cluster plate (Falcon 3047). Seedlings were treated with buffer only, mEPF2 (2.5 µM), or with Stomagen (5 µM) at room temperature before being pooled for harvest. For heat denaturation of mEPF2, the peptide solution was treated at 95 °C for 2 h before MAPK phosphorylation assays and bioassays. Plant materials were ground in liquid nitrogen, and then extracted with buffer (100 mM HEPES, pH 7.5, 5 mM EDTA, 5 mM EGTA, 2 mM DTT, 10 mM Na<sub>3</sub>VO<sub>4</sub>, 10 mM NaF, 50 mM β-glycerolphosphate, 1 mM PMSF, 1 tablet per 50 ml extraction buffer of proteinase inhibitor mixture, 10% glycerol, 7.5% (w/v) PVPP). After centrifugation at 13,000 r.p.m. for 30 min, the protein concentration was determined using a Bradford assay (Bio-Rad). Immunoblot analysis was performed using anti-phospho-ERK (1:2,000, Cell Signaling) antibody as primary antibody, and peroxidase-conjugated goat anti-rabbit IgG (1:15,000, Sigma) as secondary antibody.

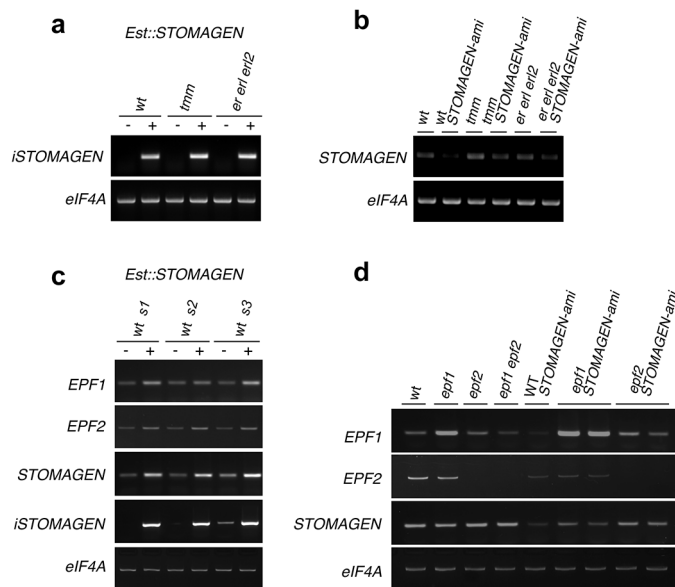
29. Shpak, E. D., Lakeman, M. B. & Torii, K. U. Dominant-negative receptor uncovers redundancy in the *Arabidopsis* ERECTA leucine-rich repeat receptor-like kinase signaling pathway that regulates organ shape. *Plant Cell* **15**, 1095–1110 (2003).
30. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
31. Pillitteri, L. J., Peterson, K. M., Horst, R. J. & Torii, K. U. Molecular profiling of stomatal meristemoids reveals new component of asymmetric cell division and commonalities among stem cell populations in *Arabidopsis*. *Plant Cell* **23**, 3260–3275 (2011).
32. Pillitteri, L. J., Sloan, D. B., Bogenschutz, N. L. & Torii, K. U. Termination of asymmetric cell division and differentiation of stomata. *Nature* **445**, 501–505 (2007).
33. Voinnet, O., Rivas, S., Mestre, P. & Baulcombe, D. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* **33**, 949–956 (2003).





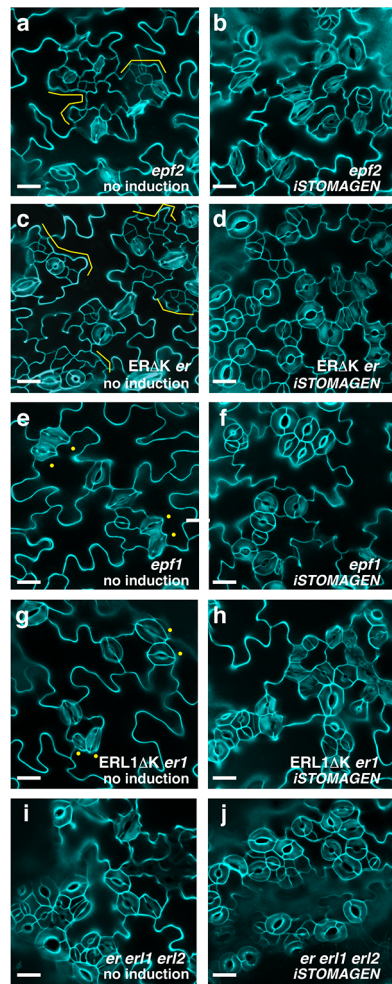
**Extended Data Figure 1 | Stomatal clustering phenotype of induced *STOMAGEN* overexpression in multiple independent transgenic lines.**

Shown are confocal microscopy images of abaxial cotyledon epidermis from 10-day-old light-grown seedlings of four independent transgenic lines carrying an oestradiol-inducible *STOMAGEN* overexpression construct (*iSTOMAGEN*). Left panels, no induction (control); right panels, oestradiol induction; each row shows representative images from individual lines. Yellow brackets indicate stomatal clusters. Images are taken under the same magnification. Scale bar, 40  $\mu\text{m}$ .  $n = 3$  for each panel.

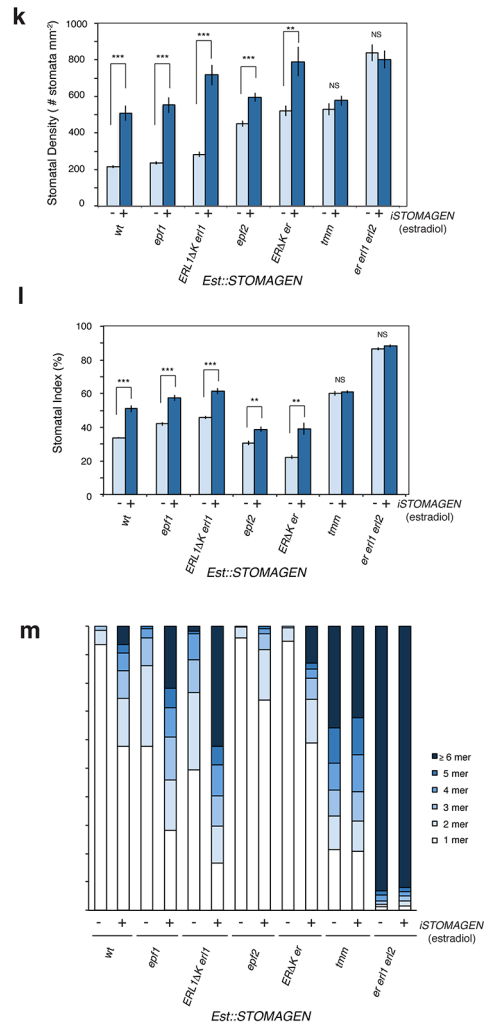


**Extended Data Figure 2 | RT-PCR analysis of *STOMAGEN* transcripts in transgenic lines used in this study.** **a**, Expression of oestradiol-inducible *STOMAGEN* transgene (*iSTOMAGEN*) in transgenic lines expressing oestradiol-inducible *STOMAGEN* overexpression (*Est::STOMAGEN*) lines from wild-type (wt), *tmm* and *erl1 erl2* triple mutant background with or without oestradiol induction. **b**, Expression of the endogenous *STOMAGEN* transcripts in each genotype carrying *STOMAGEN-ami* construct. *tmm* or *erl1 erl2* mutation does not seem to affect *STOMAGEN* transcript levels. **c**, Expression of *EPF1*, *EPF2*, total *STOMAGEN* and *STOMAGEN* transgene (*iSTOMAGEN*) transcripts in transgenic *Est::STOMAGEN* lines (in two different T1 populations (s1 and s2) and a representative T3 line (s3)) with or without oestradiol induction. *STOMAGEN* overexpression by oestradiol causes modest increase in *EPF1* and *EPF2* transcripts, which accords with increased stomatal differentiation by *iSTOMAGEN*. **d**, *EPF1*, *EPF2* and *STOMAGEN* transcript accumulation in wild-type (wt) and single- and higher-order loss-of-function mutants of *epf1*, *epf2* and *stomagen* (*STOMAGEN-ami*). For *epf1* *STOMAGEN-ami* and *epf2* *STOMAGEN-ami* lines, two different F<sub>3</sub> populations derived from the same genetic crosses were used to test the reproducibility. *STOMAGEN* expression is not influenced by *epf1* and *epf2* mutations, consistent with the proto-mesophyll expression of *STOMAGEN*. However, *EPF2* expression is reduced by *STOMAGEN-ami*, consistent with reduced stomatal cell lineages by *STOMAGEN* co-suppression. As reported, *epf1* has a T-DNA insertion within the 5' UTR<sup>3</sup>, which results in accumulation of aberrant transcripts. For all experiments, *eIF4A* was used as a control. For primer sequences see Extended Data Table 1.

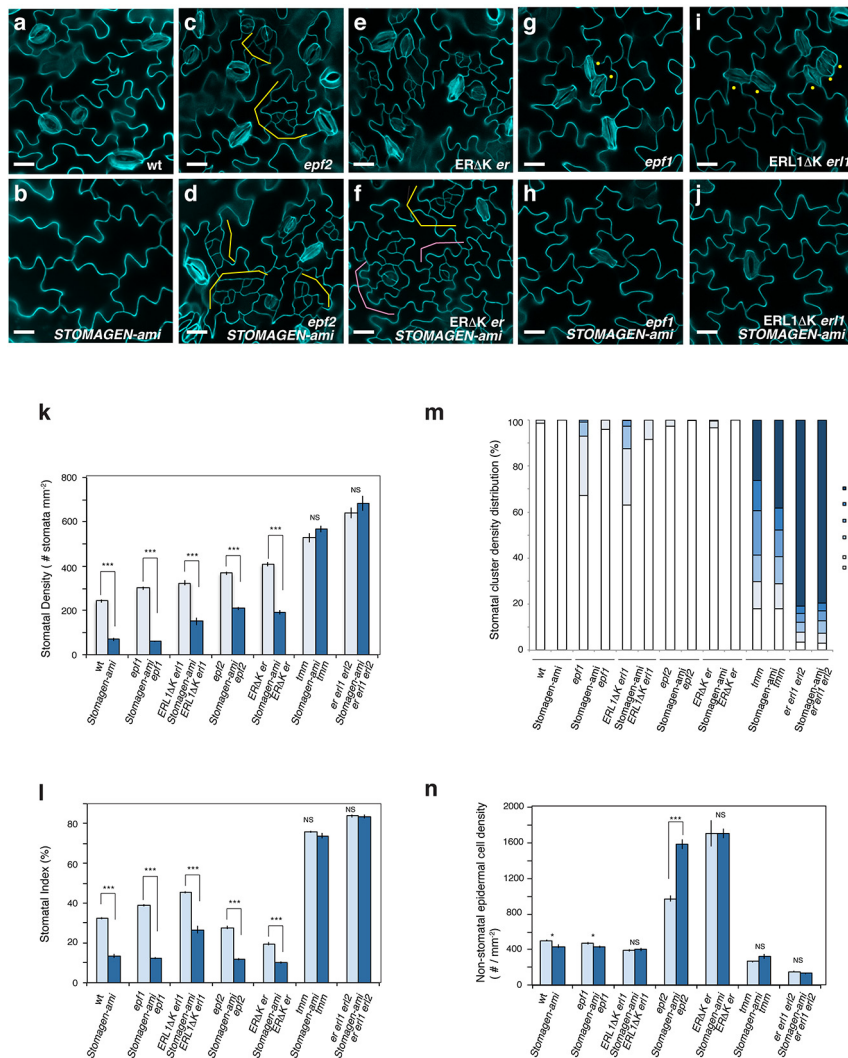




**Extended Data Figure 3 | *STOMAGEN* overexpression promotes stomatal differentiation in genetic backgrounds missing/blocking EPF2–ER and EPF1–ERL1 signalling components.** **a–j**, Representative confocal images of cotyledon abaxial epidermis from 10-day-old light-grown transgenic seedlings of the following genotypes, each carrying *Est::STOMAGEN* construct: *epf2* (**a, b**); dominant-negative ER (ER( $\Delta$ K)) in *er* (**c, d**); *epf1* (**e, f**); dominant-negative ERL1 (ERL1( $\Delta$ K)) in *erl1* (**g, h**); *er erl1 erl2* (**i, j**). For each genotype, a control uninduced phenotype (**a, c, e, g, i**) and induced *STOMAGEN* overexpression (**b, d, f, h, j**) are shown. Blocking ER or lacking EPF2 produces small stomatal-lineage cells due to excessive entry divisions (**a, c**; yellow brackets). *iSTOMAGEN* confers stomatal clusters and small stomatal-lineage cells are no longer present (**b** and **d**). Blocking ERL1 or lacking EPF1 causes a stomatal pairing due to a violation of one-cell-spacing rule (**e, g**; dots). *iSTOMAGEN* enhances stomatal cluster phenotype in these genotypes (**f, h**). *iSTOMAGEN* does not enhance stomatal clustering defects in *er erl1 erl2* (**i, j**). Images were taken under the same magnification. Scale bars, 30  $\mu$ m. *n* = 29 (**a**); *n* = 24 (**b**); *n* = 16 (**c**); *n* = 17 (**d**); *n* = 22 (**e**); *n* = 23 (**f**); *n* = 20 (**h**); *n* = 24 (**i**); *n* = 24 (**j**). **k–m**, Stomatal density



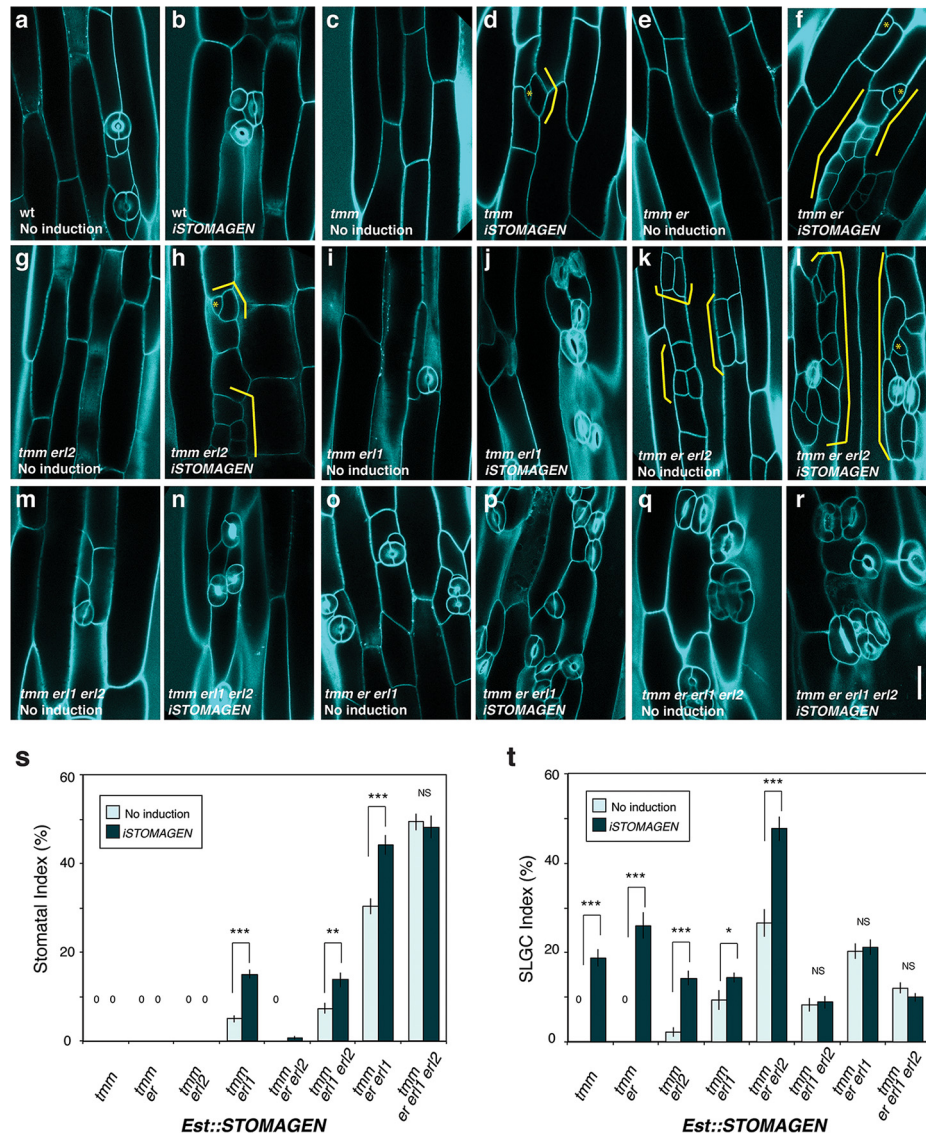
(number of stomata per mm<sup>2</sup>) (**k**); stomatal index (% of number of stomata per stomata + non-stomatal epidermal cells) (**l**); and stomatal cluster distribution (in %) (**m**) from 10-day-old abaxial cotyledons of transgenic lines of each genotype carrying *Est::STOMAGEN* construct. –, no induction; +, induced by 10  $\mu$ M oestradiol. Stomagen overexpression significantly increases stomatal density in all genotypes except for *er erl1 erl2* and *tmm*. Error bars indicate s.e.m. \*\*\**P* < 0.001; \*\**P* < 0.01; NS, not significant; Welch 2-sample *t*-test. Number of seedlings subjected to analysis, *n* = 14–16. Total numbers of stomata counted: wt, no induction 1,277, induction 2,639; *epf1*, no induction 1,390, induction 3,485; *ERL1*( $\Delta$ K) *erl1*, no induction 1,573, induction 3,991; *epf2*, no induction 2,502, induction 3,317; *ERL1*( $\Delta$ K) *er*, no induction 2,899, induction 4,397; *tmm*, no induction 2,948, induction 3,212; *er erl1 erl2*, no induction 4,454, induction 4,464. All genotypes carry *Est::STOMAGEN*. wt, no induction *n* = 16, induction *n* = 14; *epf1*, no induction *n* = 16, induction *n* = 17; *ERL1*( $\Delta$ K) *erl1*, no induction *n* = 15, induction *n* = 15; *epf2*, no induction *n* = 15, induction *n* = 15; *ERL1*( $\Delta$ K) *er*, no induction *n* = 15, induction *n* = 15; *tmm*, no induction *n* = 15, induction *n* = 15; *er erl1 erl2*, no induction *n* = 15, induction *n* = 15.



**Extended Data Figure 4 | *STOMAGEN* co-suppression results in reduced stomatal development in genetic backgrounds missing or blocked in EPF2–ER and EPF1–ERL1 signalling pathways.** **a–j**, Representative confocal images of cotyledon abaxial epidermis from 10-day-old light-grown transgenic seedlings of the following genotypes: wild type (**a**); *STOMAGEN-ami* (**b**); *epf2* (**c**); *epf2 STOMAGEN-ami* (**d**); dominant-negative ER (*ER(ΔK)*) in *er* (**e**); *ER(ΔK) er STOMAGEN-ami* (**f**); *epf1* (**g**); *epf1 STOMAGEN-ami* (**h**); dominant-negative ERL1 (*ERL1(ΔK)*) in *erl1* (**i**); *ERL1(ΔK) erl1 STOMAGEN-ami* (**j**). *STOMAGEN-ami* markedly reduces stomatal differentiation in wild type (**a**, **b**). Blocking ER or lacking EPF2 produces small stomatal-lineage cells due to excessive entry divisions (**c**, **e**; yellow brackets). *STOMAGEN-ami* exaggerates the small stomatal-lineage cells of *epf2* (**d**; yellow brackets). *STOMAGEN-ami ER(ΔK) er* shows excessive asymmetric entry as well as amplifying divisions (**f**; yellow and pink brackets, respectively). Blocking ERL1 or lacking EPF1 causes a stomatal pairing due to a violation of one-cell-spacing rule (**g**, **i**; dots). *STOMAGEN-ami* suppresses these mild stomatal pairing phenotypes and reduces stomatal differentiation (**h**, **j**). Images were taken under the same magnification. Scale bars, 30 μm. *n* = 13 (**a**); *n* = 26 (**b**); *n* = 15 (**c**); *n* = 23 (**d**); *n* = 11 (**e**); *n* = 17 (**f**); *n* = 12 (**g**); *n* = 22 (**h**); *n* = 18 (**i**); *n* = 13 (**j**).

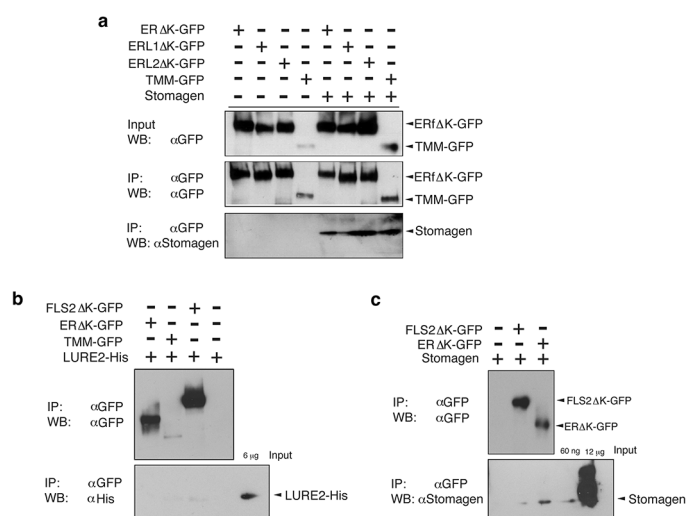
**k–n**, Stomatal density (**k**), stomatal index (**l**), stomatal cluster distribution (in %; **m**), and non-stomatal epidermal cell density (**n**) from 10-day-old abaxial cotyledons of each genotype with or without carrying *STOMAGEN-ami* construct. Error bars, s.e.m. \*\*\**P* < 0.001; \**P* ≤ 0.05; NS, not significant; Welch 2-sample *t*-test. *n* = 9–16. Total numbers of stomata counted: wt, 719; *STOMAGEN-ami*, 204; *epf1*, 1,004; *epf1 STOMAGEN-ami*, 383; *ERL1(ΔK) erl1*, 1,558; *ERL1(ΔK) erl1 STOMAGEN-ami*, 504; *epf2*, 1,505; *epf2 STOMAGEN-ami*, 1,165; *ER(ΔK) er*, 1,361; *ER(ΔK) er STOMAGEN-ami*, 782; *tmm*, 2,495; *tmm STOMAGEN-ami*, 2,688; *er erl1 erl2*, 1,853; *er erl1 erl2 STOMAGEN-ami*, 2,028. Total numbers of non-stomatal epidermal cells counted: wt, 1,494; *STOMAGEN-ami*, 1,299; *epf1*, 1,584; *epf1 STOMAGEN-ami*, 2,711; *ERL1(ΔK) erl1*, 871; *ERL1(ΔK) erl1 STOMAGEN-ami*, 1,348; *epf2*, 3,980; *epf2 STOMAGEN-ami*, 8,808; *ER(ΔK) er*, 5,739; *ER(ΔK) er STOMAGEN-ami*, 6,939; *tmm*, 790; *tmm STOMAGEN-ami*, 962; *er erl1 erl2*, 479; *er erl1 erl2 STOMAGEN-ami*, 391. wt, *n* = 8; *STOMAGEN-ami*, *n* = 8; *epf1*, *n* = 9; *epf1 STOMAGEN-ami*, *n* = 17; *ERL1(ΔK) erl1*, *n* = 13; *ERL1(ΔK) erl1 STOMAGEN-ami*, *n* = 9; *epf2*, *n* = 11; *epf2 STOMAGEN-ami*, *n* = 15; *ER(ΔK) er*, *n* = 9; *ER(ΔK) er STOMAGEN-ami*, *n* = 11; *tmm*, *n* = 8; *tmm STOMAGEN-ami*, *n* = 8; *er erl1 erl2*, *n* = 8; *er erl1 erl2 STOMAGEN-ami*, *n* = 8.





**Extended Data Figure 5 | *STOMAGEN* overexpression on stomatal development in *tmm* hypocotyl epidermis with combinatorial loss-of-function in *ER*-family genes: a complete set.** **a–r**, Representative confocal microscopy images of hypocotyl epidermis from 10-day-old light-grown transgenic seedlings of the following genotypes, each carrying *Est::STOMAGEN*: wild-type (**a**, **b**); *tmm* (**c**, **d**); *tmm er* (**e**, **f**); *tmm erl2* (**g**, **h**); *tmm erl1* (**i**, **j**); *tmm er erl2* (**k**, **l**); *tmm erl1 erl2* (**m**, **n**); *tmm er erl1* (**o**, **p**); and *tmm er erl1 erl2* (**q**, **r**). A control, uninduced phenotype (**a**, **c**, **e**, **g**, **i**, **k**, **m**, **o**, **q**); *iSTOMAGEN* results in arrested stomatal precursor cells (asterisks) and stomatal-lineage ground cells (SLGCs; brackets) in *tmm* hypocotyls (**d**). *iSTOMAGEN* triggers entry divisions in *tmm er* and *tmm erl2* (**f**, **h**; brackets), and exaggerate the SLGC clusters in *tmm er erl2* (**k**, **l**; brackets). Images were taken under the same magnification. Scale bar, 30  $\mu$ m.  $n = 19$

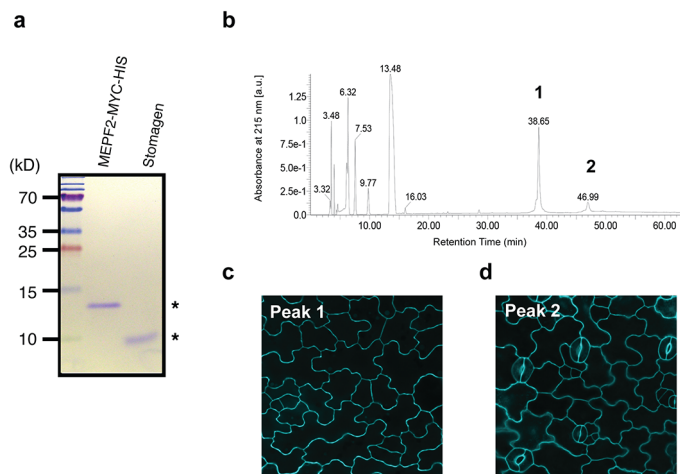
(**a**);  $n = 19$  (**b**);  $n = 20$  (**c**);  $n = 20$  (**d**);  $n = 19$  (**e**);  $n = 22$  (**f**);  $n = 20$  (**g**);  $n = 17$  (**h**);  $n = 18$  (**i**);  $n = 19$  (**j**);  $n = 19$  (**k**);  $n = 21$  (**l**);  $n = 17$  (**m**);  $n = 20$  (**n**);  $n = 19$  (**o**);  $n = 21$  (**p**);  $n = 20$  (**q**);  $n = 20$  (**r**). **s**, **t**, Stomatal index and SLGC index. **s**, \*\*\* $P < 0.0001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$  (Wilcoxon rank sum test). NS, not significant. 0, no stomata or SLGCs observed;  $n = 15$ . Total number of stomata and SLGCs counted; *tmm* non-induced, 0 and 0; induced, 0 and 211; *tmm er* non-induced, 0 and 0; induced, 0 and 308; *tmm erl2* non-induced, 0 and 32; induced, 0 and 171; *tmm erl1* non-induced, 58 and 116; induced, 142 and 138; *tmm er erl2* non-induced, 0 and 270; induced, 10 and 676; *tmm er erl1* non-induced, 422 and 283; induced, 817 and 422; *tmm erl1 erl2* non-induced, 72 and 83; induced, 163 and 97; *tmm er erl1 erl2* non-induced, 1,229 and 295; induced, 1,068 and 222.  $n = 15$  for all genotypes (**s**, **t**).



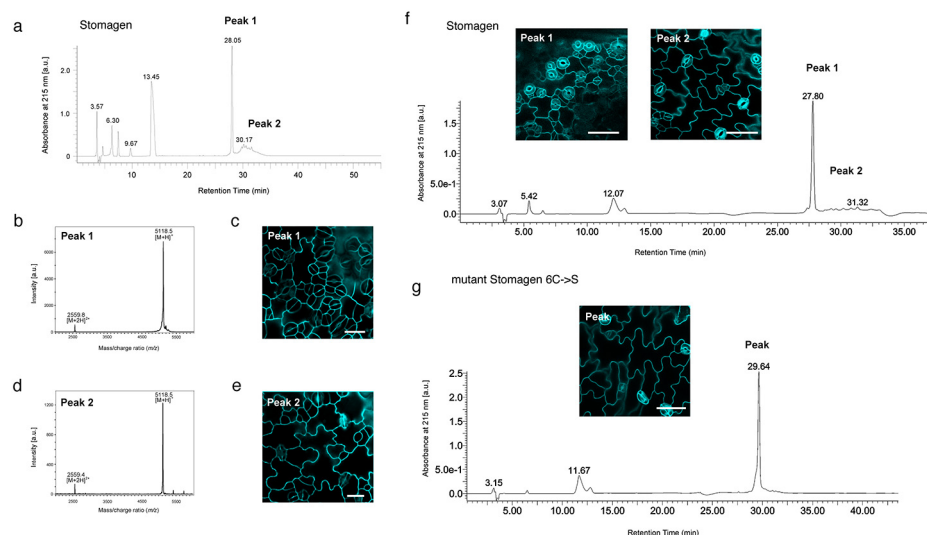
### Extended Data Figure 6 | Association of Stomagen with ER-family

**receptors and TMM.** **a**, Shown are co-immunoprecipitation assays of ligand-receptor pairs expressed in *N. benthamiana* leaves. The ectodomains and membrane-spanning domains of ER, ERL1 and ERL2 fused with GFP were separately expressed in *N. benthamiana*, and microsomal fractions were incubated with 1  $\mu$ M Stomagen peptides followed by immunoprecipitation using anti-GFP (anti-GFP) antibody. Inputs and immunoprecipitates were immunoblotted using anti-GFP (anti-GFP) or anti-Stomagen (anti-Stomagen) antibodies. Experiments were repeated three times (three biological replicates). **b**, Co-immunoprecipitation of LURE2 peptide fused with hexa-histidine tag (LURE2-His) with *N. benthamiana* microsomal fractions expressing the ectodomains and membrane-spanning domains of ER and FLS2 fused with GFP, a full-length TMM fused with GFP, or a control, uninoculated leaf sample. Immunoprecipitation was performed using anti-GFP and immunoblotted using anti-GFP (for detection of receptors) or anti-His (for detection of LURE2-His) antibodies. Experiments were repeated twice (two biological replicates). **c**, Co-immunoprecipitation of Stomagen peptide with *N. benthamiana* microsomal fractions expressing the ectodomains and membrane-spanning domains of ER and FLS2 fused with GFP or a control, uninoculated leaf sample. Immunoprecipitation was performed using anti-GFP and immunoblotted using anti-GFP (for detection of receptors) or anti-Stomagen antibodies. Experiments were repeated four times (four biological replicates).





**Extended Data Figure 7 | Purified mEPF2 and Stomagen recombinant peptides and separation of bioactive mEPF2 by reverse-phase chromatography.** **a**, SDS-PAGE gel of purified and refolded mEPF2-MYC-HIS and Stomagen recombinant peptides (asterisks). Left: molecular mass markers. **b**, HPLC chromatogram of purified, refolded mEPF2. Peaks 1 and 2 in UV chromatogram were collected and subjected to bioassays. **c**, Confocal image of cotyledon epidermis from wild-type seedling grown a solution with peak 1 for 5 days. No stoma is visible, indicating that peak 1 contains bioactive mEPF2. Scale bar, 20  $\mu$ m.  $n = 19$ . **d**, Confocal image of cotyledon epidermis from wild-type seedling grown in a solution with peak 2 for 5 days, with normal stomatal differentiation, indicating that the peptide is not bioactive. Scale bar, 20  $\mu$ m.  $n = 9$ .

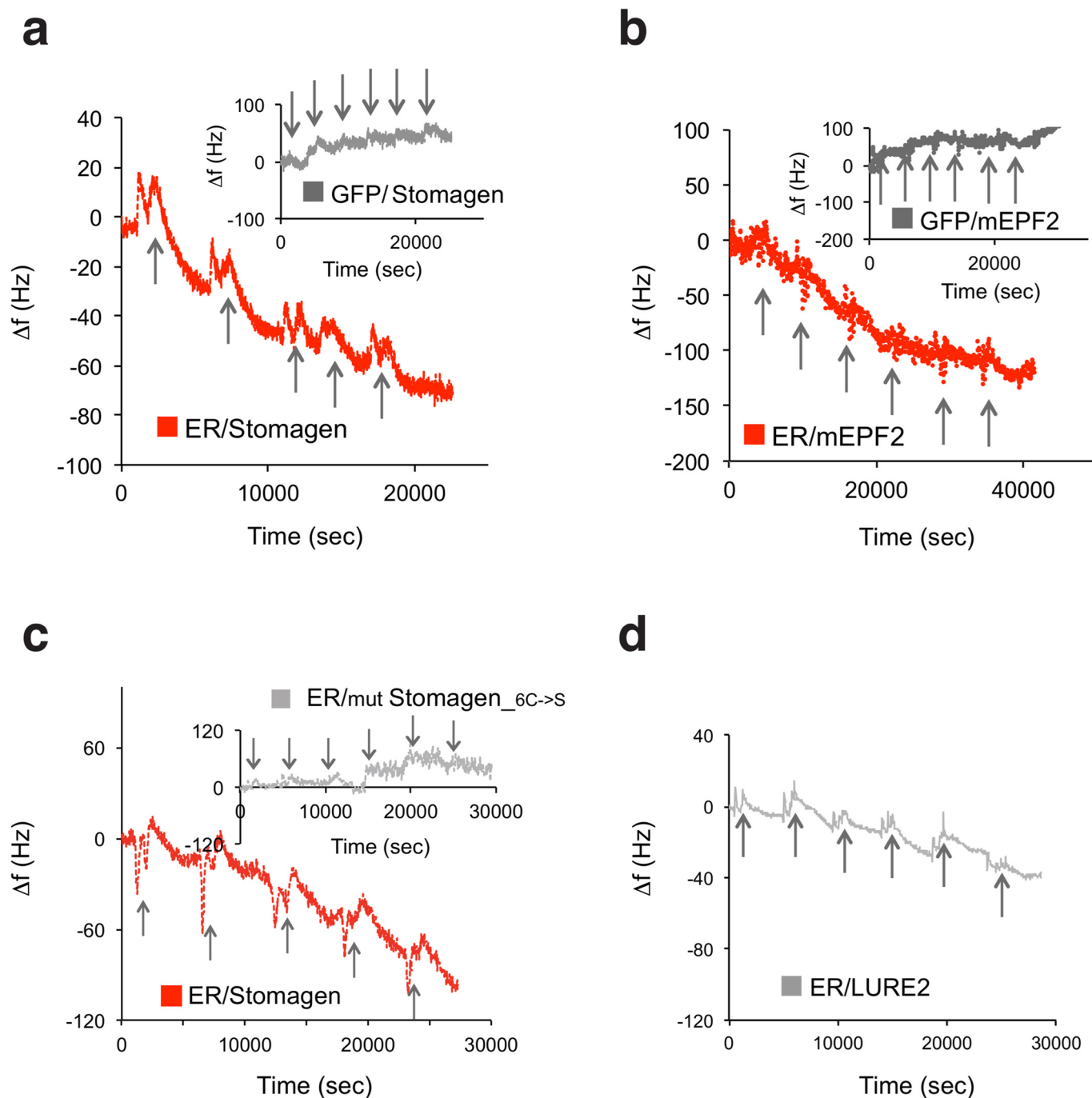


### Extended Data Figure 8 | Separation of properly folded, bioactive Stomagen and mutant Stomagen peptides by reverse-phase chromatography followed by mass spectrometry and bioassays.

**a**, HPLC chromatogram of purified, refolded Stomagen. Peaks 1 and 2 in UV chromatogram were collected and subjected to MALDI-TOF mass spectrometry (**b**, **d**) as well as for bioassays (**c**, **e**). **b**, MALDI-TOF spectrum of peak 1 from **a**. A single-charged peptide corresponding to synthetic Stomagen peptide was observed at  $m/z = 5,118.5$  ( $[M+H]^+$ ) and a double charged peptide at  $m/z = 2,559.8$  ( $[M+2H]^{2+}$ ). **c**, Confocal image of cotyledon epidermis from wild-type seedling grown in a solution with peak 1. Severe stomatal clustering and overproduction of stomata are observed. Scale bar, 20  $\mu\text{m}$ .  $n = 8$ . **d**, MALDI-TOF spectrum of peak 2 from **a**. **e**, Confocal image of cotyledon epidermis from wild-type seedling grown in a solution with peak 2 from **a**, with no stomatal clustering, indicating that the

fraction is not bioactive. Scale bar, 20  $\mu\text{m}$ .  $n = 6$ . **f**, HPLC chromatogram and bioassays of an independent batch of Stomagen peptides used for QCM analysis in direct comparison with non-folding mutant Stomagen peptides in Fig. 3c. Peaks 1 and 2 in UV chromatogram were collected and subjected for bioassays. Insets: confocal microscopy images of cotyledon epidermis from wild-type seedling grown in a solution with peak 1 (bioactive) and peak 2 (non-active) for 5 days. Scale bars, 50  $\mu\text{m}$ .  $n = 8$  (peak 1);  $n = 6$  (peak 2). **g**, HPLC chromatogram of purified, mutant Stomagen peptide in which all cysteine residues were substituted to serine residues (Stomagen\_6C->S). The mutant Stomagen peptide yielded a single peak, which was subjected for bioassays followed by confocal microscopy (inset). No stomatal clustering was observed, indicating that non-folding Stomagen peptide is not bioactive, confirming the previous results<sup>18</sup>. Scale bar, 50  $\mu\text{m}$ .  $n = 8$  for each peptide treatment.





**Extended Data Figure 9 | Raw QCM recording data.** Shown are raw recording data of frequency shifts for representative QCM analysis using biosensor chips immobilized with ER( $\Delta$ K)-GFP and GFP (**a, b, inset**) after sequential injection of active Stomagen (**a, c**), mEPF2 (**b**), non-folding, inactive mutant Stomagen (**c, inset**), or LURE2 (**d**) in increasing concentrations. Bioactive Stomagen and inactive Stomagen experiments in **c** were performed

side by side. Arrows indicate time of additional peptide application. Numbers of experiments performed for each analysis: Stomagen-ER,  $n = 3$ ; Stomagen-TMM,  $n = 2$ ; Stomagen-GFP,  $n = 3$ ; mEPF2-ER,  $n = 2$ ; mEPF2-TMM,  $n = 3$ ; mEPF2-GFP,  $n = 2$ ; inactive Stomagen<sub>C6→S</sub>-ER,  $n = 3$ ; and LURE2-ER,  $n = 2$ .

Extended Data Table 1 | List of plasmids and primers used in this study

| Plasmid ID | Description                              | Insert                       | Vector | Bac R      | Plant R |
|------------|--|------------------------------|--------|------------|---------|
| pKUT608    | STOMAGEN in pENTR                        | STOMAGEN cDNA                | pENTR  | KAN        | NA      |
| pKMP127    | proEst::STOMAGEN in pER8                 | STOMAGEN cDNA                | pER8   | SPEC/STREP | HYG     |
| pJSL92     | ERL2 genomic $\Delta$ Kinase in pENTR    | ERL2 genomic $\Delta$ Kinase | pENTR  | KAN        | NA      |
| pJSL93     | 35S::gERL2- $\Delta$ Kinase-GFP in pGWB5 | ERL2 genomic $\Delta$ Kinase | pGWB5  | KAN/HYG    | KAN/HYG |
| pJSL73     | FLS2 $\Delta$ K in pENTR                 | FLS2 $\Delta$ K cDNA no stop | pENTR  | KAN        | NA      |
| pJSL75     | 35S:FLS2 $\Delta$ K-GFP in pGWB5         | FLS2 $\Delta$ K cDNA no stop | pGWB5  | KAN/HYG    | KAN/HYG |

| Primer names            | Sequences (5' to 3')             | Purpose                     |
|-------------------------|----------------------------------|-----------------------------|
| EPFL9 1 XhoI            | CACCTCGAGATGAAGCATGAA            | molecular cloning (pKUT608) |
| EPFL9 289 SpeI rc       | ACTAGTTATCTATGACAAACAC           | molecular cloning (pKUT608) |
| FLS2 1 (GW) F           | CACCATGAAGTTACTCTCAAAGACCTTTTG   | molecular cloning (pJSL73)  |
| FLS2 2625 rc            | GATGTTGGCACTGTTGAATGAATCTGTTGC   | molecular cloning (pJSL73)  |
| FLS2 591 F              | TGTAGCAGCTGGTAACCAT              | Sequencing                  |
| <i>elF4A</i> F          | AGCCAGTGAGAATCTTGGTGAAGC         | RT-PCR                      |
| <i>elF4A</i> R          | CTAGTACGGCAGAGCAAACACAGC         | RT-PCR                      |
| <i>STOMAGEN</i> F       | TGTAGTTCAAGCCTCAAGACCTC          | RT-PCR                      |
| <i>STOMAGEN</i> R       | ACTCGTTGTACGTACAAGTTGGT          | RT-PCR                      |
| pER8 Term R             | TCGAAACCGATGATACGGACG            | RT-PCR                      |
| EPF1+207F               | ATGCCGTCTTGTGATGGTTAG            | RT-PCR                      |
| EPF1+315rc              | TCAAGGGACAGGGTAGGACTT            | RT-PCR                      |
| EPF2.1.cDNA.xhoI        | CACCCTCGAGATGACGAAGTTTGTACGCAAGT | RT-PCR                      |
| EPF2.360.cDNA.ecoRI.rc2 | CGGAATTCTAGCTCTAGATGGCACGTGATAG  | RT-PCR                      |



# HIF-driven *SF3B1* induces KHK-C to enforce fructolysis and heart disease

Peter Mirtschink<sup>1\*</sup>, Jaya Krishnan<sup>1\*†</sup>, Fiona Grimm<sup>1†</sup>, Alexandre Sarre<sup>2</sup>, Manuel Hörl<sup>3</sup>, Melis Kayikci<sup>4†</sup>, Niklaus Fankhauser<sup>1</sup>, Yann Christinat<sup>1</sup>, Cédric Cortijo<sup>1</sup>, Owen Feehan<sup>1</sup>, Ana Vukolic<sup>1</sup>, Samuel Sossalla<sup>5</sup>, Sebastian N. Stehr<sup>6</sup>, Jernej Ule<sup>4†</sup>, Nicola Zamboni<sup>3</sup>, Thierry Pedrazzini<sup>2</sup> & Wilhelm Krek<sup>1</sup>

**Fructose is a major component of dietary sugar and its overconsumption exacerbates key pathological features of metabolic syndrome. The central fructose-metabolising enzyme is ketohexokinase (KHK), which exists in two isoforms: KHK-A and KHK-C, generated through mutually exclusive alternative splicing of *KHK* pre-mRNAs. KHK-C displays superior affinity for fructose compared with KHK-A and is produced primarily in the liver, thus restricting fructose metabolism almost exclusively to this organ. Here we show that myocardial hypoxia actuates fructose metabolism in human and mouse models of pathological cardiac hypertrophy through hypoxia-inducible factor 1 $\alpha$  (HIF1 $\alpha$ ) activation of *SF3B1* and *SF3B1*-mediated splice switching of KHK-A to KHK-C. Heart-specific depletion of *SF3B1* or genetic ablation of *Khk*, but not *Khk*-A alone, in mice, suppresses pathological stress-induced fructose metabolism, growth and contractile dysfunction, thus defining signalling components and molecular underpinnings of a fructose metabolism regulatory system crucial for pathological growth.**

Heart failure is a leading cause of death worldwide. It is often preceded by hypertrophic growth of the organ as it strains to maintain function under conditions of pathologic stress<sup>1</sup> and is associated with an increased demand for oxygen, owing to changes in coronary resistance, cardiac vascularisation and perfusion<sup>2</sup>. The ensuing state of hypoxia activates hypoxia-inducible factor (HIF)1 $\alpha$  that mediates adaptive transcriptional responses to low-oxygen tension<sup>3</sup> to support cardiomyocyte growth, myocardial angiogenesis and reprogramming of metabolism<sup>4</sup>. Unremitted pathologic stress leads ultimately to functional deterioration of the heart with the clinical signs and symptoms of heart failure<sup>5</sup>.

Given that both HIF1 $\alpha$  transcription and alternative pre-mRNA splicing are central to the control of metabolism, we embarked on a systematic and functional analysis of HIF1 $\alpha$ -dependent splice factor expression and pre-mRNA splicing in pathologic cardiac hypertrophy. We identified splice factors as HIF1 $\alpha$  targets, which prompted the analysis of their RNA targets, unveiling mechanistic and functional linkages between HIF1 $\alpha$ , splice factor 3b subunit 1 (*SF3B1*), KHK-C splice isoform production and fructose metabolism in cardiac hypertrophy.

## The HIF1 $\alpha$ -*SF3B1*-KHK-C axis

To deduce changes in both gene expression and alternative splicing as a function of Hif1 $\alpha$ , we followed the experimental workflow illustrated in Extended Data Fig. 1a. RNA isolated from normoxic and hypoxic neonatal mouse cardiomyocytes (NMCs), depleted (or not) of Hif1 $\alpha$  by specific short hairpin RNAs (shHif1 $\alpha$ ), was subjected to a splice-junction microarray analysis. Genes exhibiting a Hif1 $\alpha$ -dependent expression profile were manually checked for an involvement in pre-mRNA splicing and, in parallel, matched to a list of

329 unique genes comprised of validated and *in silico*-predicted splice factors<sup>6–8</sup> (Supplementary Fig. 1 and Supplementary Table 1). Twenty genes with characteristics of splice factors/regulators displayed a Hif1 $\alpha$ -dependent expression pattern (Fig. 1a), of which four, *C1qbp*, *Hnrnp3*, *Imjd6* and *Sf3b1*, contained conserved hypoxia-responsive elements (HREs) in their promoters (Extended Data Fig. 1b).

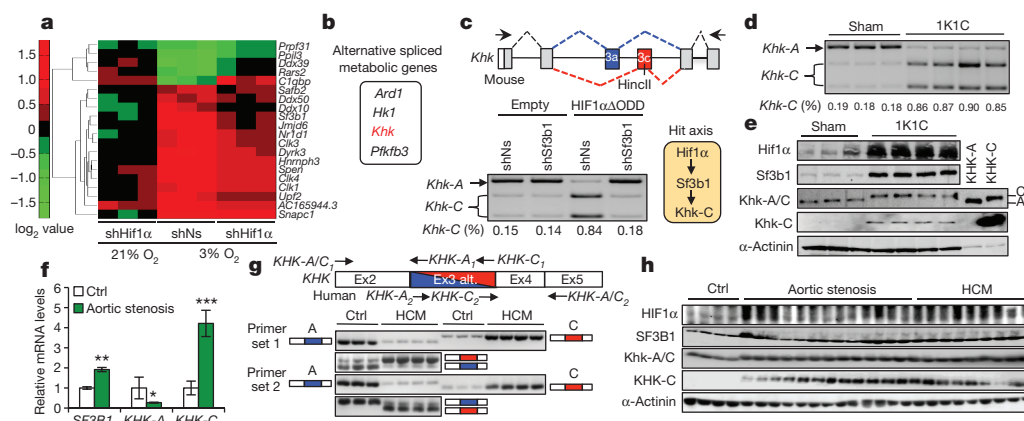
Data derived from the splice-junction microarray analysis were also interrogated for changes in gene splicing as a function of Hif1 $\alpha$  and further filtered for their potential association to metabolism, leading to *Ard1* (also known as *Naa10*), *Hk1*, *Khk*, and *Pfkfb3*, which exhibited the most consistent and significant changes (Fig. 1b). The Hif1 $\alpha$  dependency of *Khk* alternative splicing implied a link between hypoxia and fructose metabolism, especially since the dominant *Khk* isoform induced by hypoxia, *Khk*-C, has a superior activity over *Khk*-A<sup>9</sup> (Supplementary Fig. 2).

Analysis of *Khk* alternative splicing in NMCs by real-time PCR (RT-PCR) revealed Hif1 $\alpha$ -mediated, *Sf3b1*-dependent induction of *Khk*-C mRNAs, while mRNA levels of *Khk*-A declined (Fig. 1c and Extended Data Fig. 1c–e). Isoproterenol treatment yielded similar results (Extended Data Fig. 1f, g). Also, in mouse models of cardiac hypertrophy, specifically one-kidney one-clip (1K1C), trans-aortic constriction (TAC) and isoproterenol infusion, *Khk*-C mRNA and protein production increased over sham-treated control mice (Fig. 1d, e and Extended Data Fig. 1h–n). Hif1 $\alpha$  target and hypertrophic marker gene expression were dramatically increased in all three cardiomyopathy models (Extended Data Fig. 1o and Supplementary Fig. 3a–c).

Analysis of HIF1 $\alpha$ , *SF3B1* and KHK-C in ventricular biopsies of aortic stenosis or hypertrophic cardiomyopathy (HCM) patients corroborated the results obtained in mice both at the level of mRNA and protein expression (Fig. 1f–h and Extended Data Fig. 1p). mRNAs of

<sup>1</sup>Institute of Molecular Health Sciences, ETH Zurich, 8093 Zürich, Switzerland. <sup>2</sup>Department of Medicine, University of Lausanne, 1011 Lausanne, Switzerland. <sup>3</sup>Institute of Molecular Systems Biology, ETH Zurich, 8093 Zürich, Switzerland. <sup>4</sup>MRC-Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. <sup>5</sup>Universitätsmedizin Göttingen, Klinik für Kardiologie und Pneumologie, D-37075 Göttingen, and DZHK (German Centre for Cardiovascular Research), Partner Site Göttingen, Germany. <sup>6</sup>Department of Anesthesiology and Critical Care Medicine, University Hospital Jena, 07747 Jena, Germany. <sup>†</sup>Present addresses: MRC Clinical Sciences Centre London, Imperial College London, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK (J.K.); MRC National Institute for Medical Research, The Ridgeway, London NW7 1AA, UK (F.G.); MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK (M.K.); Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK (J.U.).

\*These authors contributed equally to this work.



**Figure 1 | Identification of Hif1 $\alpha$ -dependent splicing factors and splicing events.** **a**, Expression of splice factors in NMCs transduced as indicated versus NMCs treated with non-silencing shRNA (shNs) and cultured at 21% O<sub>2</sub>. **b**, Hit list of metabolic genes. **c**, **d**, NMCs transduced as indicated (**c**) and ventricular samples from sham- and 1K1C-treated mice (**d**) were analysed for *Khk*-A and *Khk*-C expression by RT-PCR and HincII digestion. Primers to amplify the *KHK* variable region are indicated (**c**, upper). The hypothetical HIF1 $\alpha$ -SF3B1-KHK-C axis (**c**, right). **e**, Ventricular samples from sham- and 1K1C-treated mice and lysates of NMCs overexpressing KHK-A or KHK-C were processed for immunoblotting with antibodies against denoted proteins. A and C indicate the differential migration of KHK-C and KHK-A on SDS gels. **f**, Biopsies of left ventricles from aortic stenosis patients and healthy

controls were assessed for *SF3B1*, *KHK*-A and *KHK*-C expression by qPCR ( $n = 16$  for patient samples and  $n = 6$  for controls; error bars are s.e.m.;  $**P < 0.01$ ;  $***P < 0.001$ ; two-tailed unpaired  $t$ -test). **g**, Top, schematic representation of exon 2 to 5 of human *KHK*. Alternative usage of exon 3 (Ex3 alt.) is displayed in blue (exon 3A) and red (exon 3C). Primer sets location and orientation used for RT-PCR analysis, are indicated with arrows. Bottom, assessment of *KHK*-A and *KHK*-C expression by RT-PCR in ventricles of controls and HCM patients using primer set 1 and 2 as described above. **h**, Lysates of left-ventricular samples from healthy controls and aortic stenosis or HCM patients were processed for immunoblotting with antibodies against denoted proteins.

HIF1 $\alpha$  targets and hypertrophic markers were also induced (Extended Data Fig. 1q and Supplementary Fig. 3d). Thus, HIF1 $\alpha$ , SF3B1 and KHK-C are coordinately upregulated in different mouse models of cardiac hypertrophy and biopsies of human cardiomyopathy patients.

### The HIF1 $\alpha$ target SF3B1 mediates splicing of KHK

Sf3b1 appears to represent a direct Hif1 $\alpha$  target. First, treatment of NMCs with phenylephrine or isoproterenol, known inducers of Hif1 $\alpha$  and Hif1 $\alpha$ -target genes<sup>10</sup>, increased *Sf3b1* mRNA expression in a Hif1 $\alpha$ -dependent manner (Extended Data Fig. 2a). Second, *Sf3b1* mRNA expression also increased upon depletion of the Von Hippel-Lindau protein (pVhl) or ectopic expression of HIF1 $\alpha$  lacking the oxygen-dependent degradation domain (HIF1 $\alpha$ ΔODD) (Extended Data Fig. 2b). Third, the promoter of SF3B1 contains a highly conserved HRE (Extended Data Fig. 2c). Fourth, expression of HIF1 $\alpha$ ΔODD activated a mouse wild-type *Sf3b1*-promoter-luciferase reporter but not the corresponding reporter harbouring a mutant HRE (Fig. 2a). Fifth, chromatin immunoprecipitation (ChIP) assays demonstrate an association of Hif1 $\alpha$  with the HRE in the *Sf3b1* promoter in native chromatin (Fig. 2b).

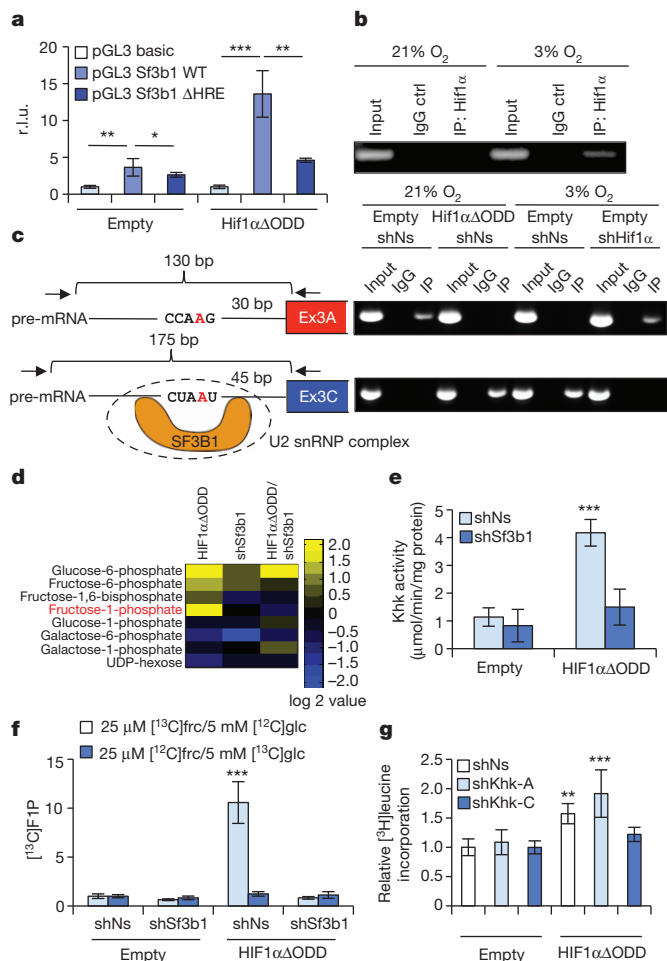
Next we measured the incorporation of radioactively labelled [<sup>3</sup>H]leucine in NMCs depleted of Sf3b1 and exposed to various treatments including phenylephrine or isoproterenol (Extended Data Fig. 2d), expression of HIF1 $\alpha$ ΔODD (Extended Data Fig. 2e), hypoxia (Extended Data Fig. 2f) or simultaneous depletion of pVhl (Extended Data Fig. 2g). We observed that [<sup>3</sup>H]leucine incorporation (Extended Data Fig. 2i–k),  $\alpha/\beta$  myosin heavy chain mRNA ratio (Extended Data Fig. 2l–o) and cell size (Extended Data Fig. 2p) changed as a function of Sf3b1. Other components of the U2 small nuclear ribonucleoprotein (U2 snRNP) complex, namely Sf3b2, Sf3b3 and Sf3a3, neither increased in abundance in response to Hif1 $\alpha$  expression nor did they affect, when downregulated by different shRNAs, to a measurable extent, Hif1 $\alpha$ -stimulated cell growth (Extended Data Fig. 2q, r and Supplementary Fig. 4).

Also, known parameters of Hif1 $\alpha$ -driven metabolism including extracellular acidification rate (ECAR; an indicator lactate production) and oxygen consumption rate (OCR; an indicator of oxidative phosphorylation) changed in a Sf3b1-dependent manner in NMCs.

At basal levels, ECAR was not affected by depletion of Sf3b1 but increased upon addition of isoproterenol (Extended Data Fig. 2s, t). OCR was repressed, as expected, in isoproterenol-treated cells upon addition of either oleic acid or carbonyl cyanide-4-(trifluoromethoxy)-phenylhydrazone, an effect diminished by depletion of Sf3b1 (Extended Data Fig. 2u). Ectopic expression of HIF1 $\alpha$ ΔODD produced similar effects on ECAR and OCR as isoproterenol (Extended Data Fig. 2v–z), while depletion of Sf3b1 alone had no effect (Supplementary Fig. 5). In sum, these results suggest a key contributory role for Sf3b1 in Hif1 $\alpha$ -mediated growth and metabolic programs.

Sf3b1 participates in the catalysis of splicing<sup>11</sup>, can be crosslinked to the 5' and 3' of the branch point adenosine<sup>12</sup>, and displays differential effects on specific splice targets under conditions of reduced function<sup>13,14</sup>. Indeed, splice-junction analysis revealed widespread effects on both alternative splicing and gene expression that included changes in *Khk* isoform expression when Sf3b1 was depleted in hypoxic NMCs (Supplementary Tables 2 and 3). RNA-immunoprecipitations of Sf3b1 from extracts of NMCs exposed to hypoxia or HIF1 $\alpha$ ΔODD overexpression, demonstrate that Sf3b1 occupies RNA sequences upstream of exon 3C (and not of exon 3A) of *Khk* pre-mRNA in hypoxia but not normoxia (Fig. 2c). Upon depletion of Hif1 $\alpha$ , Sf3b1 was redirected towards branch-point sequences of exon 3A (Fig. 2c). This suggests that Sf3b1 association with RNA sequences upstream of exon 3C of *Khk* requires Hif1 $\alpha$ -signalling.

Sf3b1 acts as a direct mediator of *Khk* splicing as its overexpression caused a shift from *Khk*-A to *Khk*-C (Extended Data Fig. 3a). Selective depletion of endogenously (but not exogenously) produced Sf3b1 prevented this shift (Extended Data Fig. 3a). Quantitative PCR (qPCR) analysis confirmed the overexpression and/or reduction of Sf3b1 mRNAs where expected (Extended Data Fig. 3b). Concomitant with the shift in *Khk* isoform expression towards *Khk*-C, cell growth was induced (Extended Data Fig. 3c). Neither the overproduction of SF3A3 or SF3B3 (Extended Data Fig. 3d, e), nor the depletion of Sf3a3, Sf3b3 or Sf3b2 by corresponding shRNAs, generated a similar robust splice shift towards *Khk*-C despite efficient depletion of the respective proteins (Extended Data Fig. 3f, g and Supplementary Fig. 4). Cell growth was also not affected by SF3A3 or SF3B3 expression (Extended Data Fig. 3h).



**Figure 2 | *SF3B1* is a HIF1 $\alpha$ -target gene that mediates splicing of KHK.**

**a**, Co-transfection of wild type (WT) or HRE-mutated *SF3B1* promoters fused to luciferase with either an empty vector control or HIF1 $\alpha$ ΔODD (values are relative light units (r.l.u.);  $n = 4$  biological replicates per group). **b**, NMCs were cultured as indicated and processed for chromatin immunoprecipitation with a HIF1 $\alpha$ -specific antibody (IP: HIF1 $\alpha$ ) or with a control isotype-matched antibody (IgG control). **c**, Schematic representing branch-point sequences located in front of exon 3A (upper left) or exon 3C (lower left). Branch point adenine is labelled in red. Location and orientation of primers (arrows), expected length of fragments to be amplified by RT-PCR and potential Sf3b1 binding sites are depicted. NMCs treated as indicated were processed for RNA-immunoprecipitation assay with antibodies specific for Sf3b1 (IP) or control-isotype matched antibody (IgG). Recovered RNA was purified and analysed by RT-PCR (upper right and lower right). **d**, Mass spectrometry-based metabolomic profiling of NMCs transfected with lentiviruses as denoted focused on hexose-phosphates. Metabolite values are compared to amounts in control transduced NMCs ( $n = 4$  biological replicates per group). **e**, KHK activity in lysates of NMCs transfected with lentiviruses as indicated ( $n = 7$  biological replicates per group). **f**, Relative amount of [ $^{13}\text{C}$ ]fructose-1-phosphate derived from [ $^{13}\text{C}$ ]fructose ([ $^{13}\text{C}$ ]frc) or [ $^{13}\text{C}$ ]glucose ([ $^{13}\text{C}$ ]glc) in NMCs transfected as denoted ( $n = 4$  biological replicates per group). **g**, [ $^3\text{H}$ ]leucine incorporation in NMCs infected as indicated ( $n = 3$  biological replicates per group, data show 1 of 3 representative experiments). For panels **a**, **e–g**, \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Two-tailed unpaired  $t$ -test (**a**) one-way ANOVA followed by Dunnett's multiple comparison post-test (**e–g**). Error bars are s.d. (**a**, **g**) or s.e.m. (**e**, **f**).

Next we proceeded with mass-spectrometry-based metabolomic profiling of hexose-phosphates in NMCs expressing HIF1 $\alpha$ ΔODD in the absence or presence of shSf3b1. We observed profoundly increased levels of fructose-1-phosphate (F1P) and glucose-6-phosphate in response to HIF1 $\alpha$  activation (Fig. 2d), but only the level of the former was Sf3b1-dependent (Fig. 2d). In accord with increased levels of F1P, HIF1 $\alpha$ ΔODD also increased total cellular

Khk activity by 3.6-fold compared to control NMCs (Fig. 2e). In a comparative analysis, overexpressed human KHK-C displayed, in NMCs depleted for endogenous Khk, a 3.7-fold higher kinase activity than KHK-A (Extended Data Fig. 3i, j). Thus, activation of the Hif1 $\alpha$ -Sf3b1 axis activates Khk-C to stimulate F1P production.

To assess whether increased levels of F1P are derived from external fructose or from conversion of glucose via the polyol pathway<sup>15</sup>, we supplied NMC culture media with [ $^{13}\text{C}$ ]fructose and unlabelled glucose or [ $^{13}\text{C}$ ]glucose and unlabelled fructose and measured [ $^{13}\text{C}$ ]F1P production. [ $^{13}\text{C}$ ]F1P levels increased upon HIF1 $\alpha$ ΔODD expression selectively in NMCs cultured with [ $^{13}\text{C}$ ]fructose and unlabelled glucose but not [ $^{13}\text{C}$ ]glucose and unlabelled fructose in a manner dependent on Sf3b1 (Fig. 2f). Moreover, KHK-C, but not KHK-A, expression induced accumulation of [ $^{13}\text{C}$ ]F1P levels in a time-dependent fashion in NMCs cultured with [ $^{13}\text{C}$ ]fructose (Extended Data Fig. 3k). *Glut5* fructose transporter mRNA expression was likewise induced by KHK-C and HIF1 $\alpha$ ΔODD (Extended Data Fig. 3l).

Khk-C is also a critical downstream mediator of Hif1 $\alpha$ -induced cell growth as selective depletion of Khk-C, but not Khk-A, rescued HIF1 $\alpha$ ΔODD-induced cardiomyocyte growth (Fig. 2g and Extended Data Fig. 3m). Knockdown efficiency and isoform specificity of the corresponding shRNAs targeting Khk isoforms was confirmed by qPCR (Extended Data Fig. 3n). Taken together, Sf3b1 appears to be a direct mediator of Khk pre-mRNA splicing, resulting in Khk-C expression, increased Khk activity, fructose uptake and cell growth.

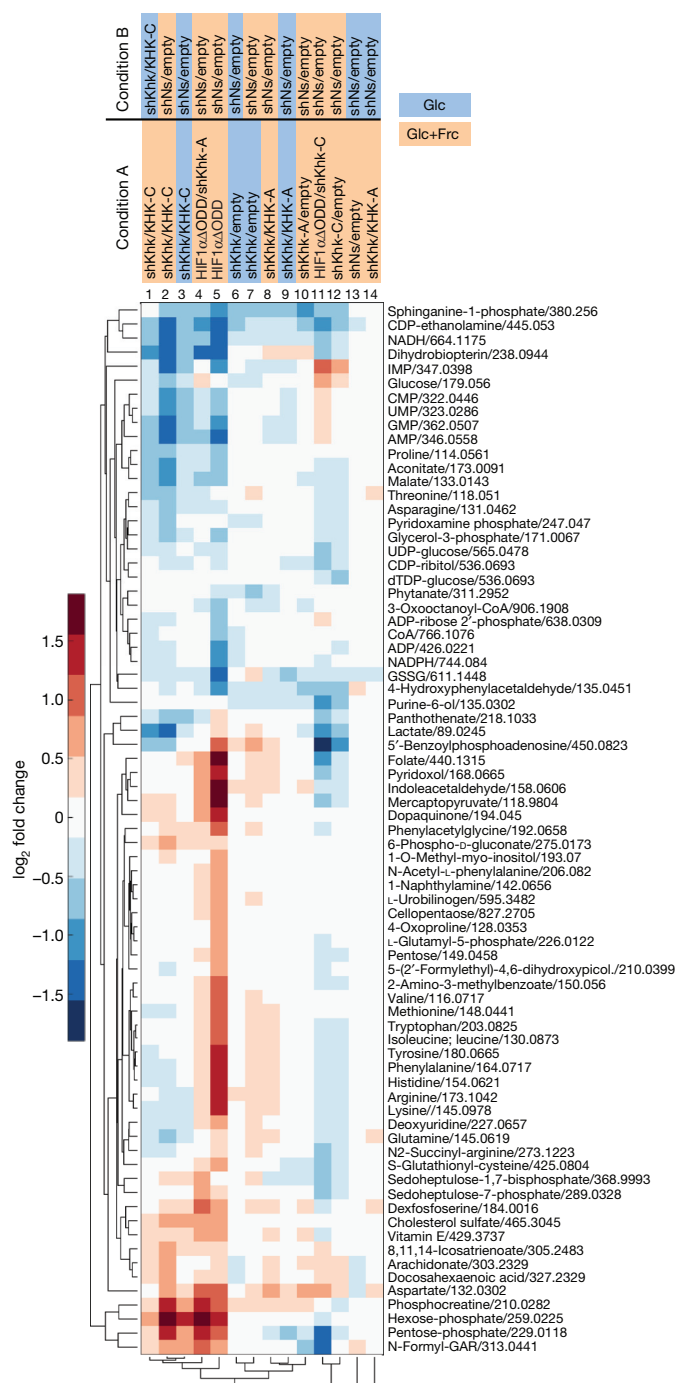
## Expression of KHK-C promotes anabolic metabolism

To characterize metabolic phenotypes in NMCs associated with activation of the HIF1 $\alpha$ -Sf3B1-KHK-C axis and as a function of fructose availability, we performed metabolomic profiling. These efforts established the following experimental points. First, in the presence of fructose and glucose, HIF1 $\alpha$ ΔODD expression increased glycolysis (for example, hexose-phosphates), amino acid metabolism and biosynthesis (for example, dextroserine, aspartate, arginine), the non-oxidative pentose-phosphate pathway (PPP; for example, sedoheptulose-7-phosphate, pentose phosphates) and to a lesser degree also the oxidative PPP (for example, 6-phospho-D-gluconate, pentose phosphates) and intermediates of DNA/RNA biosynthesis (for example, N-formyl-GAR, folate, deoxyuridine) (Fig. 3). Changes in redox balance are exemplified by lower levels of NADPH, NADH and of the reduced state of glutathione (GSSG) (Fig. 3, lanes 4 and 5). Reduced levels of these metabolites are probably the result of extensive utilization of reducing equivalents for macromolecular biosynthesis<sup>16</sup>. Second, depletion of Khk-C (but not Khk-A) reverted HIF1 $\alpha$ ΔODD-induced metabolic changes (Fig. 3, compare lane 11 with 5 and 4), while its overexpression mimicked them (Fig. 3, compare lane 1–3 and 4–5). Third, when KHK-C was overexpressed under fructose-lacking conditions, levels of glycolytic and PPP metabolites decreased and the reduction of NADPH and NADH was less pronounced (Fig. 3, compare lane 2 and 3; Extended Data Fig. 4a). Hence, exogenous fructose contributes significantly to the metabolic effects mediated by Khk-C.

In accord with the above, we detected [ $^3\text{H}$ ] radiolabel derived from [ $^3\text{H}$ ]fructose in the RNA, DNA and protein fractions of NMCs in a HIF1 $\alpha$ ΔODD- and KHK-C-dependent manner (Extended Data Fig. 4b). Expression of KHK-C stimulated also glucose uptake in the presence of fructose (Extended Data Fig. 4c). Moreover, [ $^3\text{H}$ ]leucine incorporation into NMCs increased as a function of fructose availability in a KHK-C-dependent fashion (Extended Data Fig. 4d). Increasing glucose availability in the presence of fructose had similar effects (Extended Data Fig. 4d). Immunoblotting confirmed the expected expression pattern of KHK-A and KHK-C (Supplementary Fig. 6a). These results suggest a critical role for KHK-C in HIF1 $\alpha$ -dependent metabolic reprogramming and growth.

Since phosphorylation of fructose by Khk-C is unrestrained, it could contribute to ATP-depletion as shown previously in liver<sup>17</sup>,





**Figure 3 | Anabolic metabolism and growth as a function of KHK-C.** Untargeted analysis of metabolites in NMCs treated as denoted. Depicted are metabolites with log<sub>2</sub>(fold change) > 0.5 and adjusted *P* value < 0.01 in at least one treatment group compared to corresponding control (*n* = 4 biological replicates per group).

thereby limiting ATP-driven allosteric inactivation of glycolysis at the level of phosphofructokinase (PFK)<sup>18,19</sup>. When HIF1αΔODD was overexpressed in NMCs, the ADP/ATP ratios increased (Extended Data Fig. 4e). Simultaneous depletion of either Sf3b1 or Khk-C (but not Khk-A) prevented this (Extended Data Fig. 4e). Khk-C but not Khk-A, when overexpressed, enhanced ADP/ATP ratios (Extended Data Fig. 4e). Thus, activation of the HIF1α-SF3B1-KHK-C axis increases ADP/ATP ratios.

Hypoxia or activation of HIF1α drives an accumulation of lipids in heart hypertrophy<sup>10,20</sup>. Since the F1P product DHAP is important for

formation of the glycerol backbone in triglyceride synthesis, we tested the Khk-C-dependency of this phenomenon by Oil Red O staining of lipid droplets in oleic-acid-treated NMCs. Overexpression of HIF1αΔODD or exposure to hypoxia induced lipid droplet formation, which was prevented upon depletion of Khk-C but not Khk-A (Extended Data Fig. 4f–h). Similar results were obtained when lipid levels were quantified with Nile Red (Supplementary Fig. 6b, c). Khk-C, but not Khk-A, was also required for phenylephrine-induced lipid accumulation (Supplementary Fig. 6d). Fructose carbon contributes to triglyceride synthesis in NMCs as measured by the amount of fructose-derived [<sup>3</sup>H] in the lipid fraction of NMCs expressing HIF1αΔODD (Supplementary Fig. 6e). Downregulation of Khk-A had no effect, but depletion of either Khk-C, Glut5 or Sf3b1 prevented HIF1α-driven conversion of fructose carbon into lipids (Supplementary Fig. 6e, f). In a gain-of-function experiment, overexpression of KHK-C promoted fructose carbon to lipid conversion while KHK-A expression did not (Supplementary Fig. 6e).

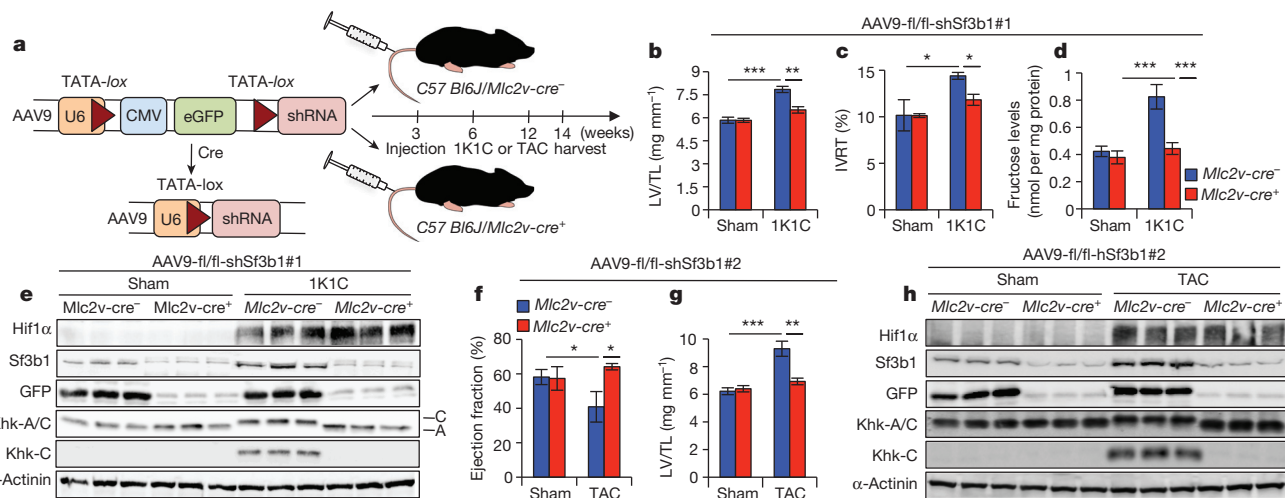
We also investigated the contribution of the two different Khk-isoforms to Hif1α-driven changes in OCR and ECAR in NMCs (Extended Data Fig. 4i–l). Khk-A depletion affected neither the ECAR nor the OCR of HIF1αΔODD-expressing cells (Extended Data Fig. 4i), while depletion of Khk-C did (Extended Data Fig. 4j). In the absence of HIF1αΔODD, neither Khk-A nor Khk-C depletion alone demonstrated any measurable effects on ECAR or OCR (Supplementary Fig. 6g, h). In a converse experiment, overexpression of KHK-A alone had little, if any, effect on these parameters while KHK-C overexpression increased ECAR and decreased OCR (Extended Data Fig. 4k, l). Thus, by stimulating fructose metabolism in cardiomyocytes, Khk-C expression promotes conversion of fructose carbon to lipids, suppresses mitochondrial oxidative phosphorylation and increases glycolysis.

### Cardiac growth and dysfunction depends on Sf3b1

Next we assessed whether myocardial fructose and glucose uptake is induced in an *in vivo* model of cardiac hypertrophy (Extended Data Fig. 5a). When TAC-treated and sham-operated control mice were simultaneously exposed to [<sup>3</sup>H]fructose and [<sup>14</sup>C]deoxyglucose, they accumulated to higher levels in diseased hearts (Extended Data Fig. 5b, c). Hearts of mice after 1K1C surgery (Extended Data Fig. 5d), demonstrated a 2.5-fold increase of fructose levels (Extended Data Fig. 5e). [<sup>14</sup>C]fructose or [<sup>14</sup>C](deoxy)glucose uptake was also stimulated under conditions of hypertrophy in *ex vivo* ventricular tissue (Extended Data Fig. 5f, g). In line with this, *Glut5* expression was induced in these hypertrophy models (Extended Data Fig. 5h, i) in hearts of isoproterenol-treated mice and in left-ventricular biopsies of aortic stenosis and HCM-patients (Extended Data Fig. 5j–l). In contrast, expression of aldose reductase (*AKR1B1*), which converts glucose to sorbitol in the polyol pathway, was not induced in the latter samples (Extended Data Fig. 5k, l). Thus, cardiac hypertrophy is associated with increased fructose levels and uptake.

To study the Sf3b1-dependency of these phenomena in response to pathologic stress, we designed a viral-based vector system enabling ventricular-specific knockdown of Sf3b1 that is based on an AAV9 backbone containing a U6 promoter followed by a TATA-*lox* site, a CMV promoter driving eGFP, another TATA-*lox* site and shRNAs targeting Sf3b1 (Fig. 4a). We constructed two viruses each expressing different shRNAs targeting Sf3b1 (AAV9-fl/fl-shSf3b1#1 and AAV9-fl/fl-shSf3b1#2). Validation of these recombinant viruses was performed both *in vitro* in NMCs and *in vivo* in *Mlc2v-cre*<sup>+</sup> and *Mlc2v-cre*<sup>−</sup> mice (Extended Data Fig. 5m, n and Supplementary Fig. 7a–d).

While injection of AAV9-fl/fl-shSf3b1#1 into *Mlc2v-cre*<sup>+</sup> or *Mlc2v-cre*<sup>−</sup> mice failed to cause any effects on ventricular size in sham-operated mice, AAV9-fl/fl-shSf3b1#1 injection protected *Mlc2v-cre*<sup>+</sup> but not *Mlc2v-cre*<sup>−</sup> littermates from 1K1C-induced cardiac hypertrophy (Fig. 4b) and diastolic dysfunction (Fig. 4c and Extended Data Table 1).



**Figure 4 | SF3B1 is required for pathologic cardiac growth and function.** **a**, Schematic representation of the AAV9-fl/shSf3b1 virus before and after Cre-mediated recombination (left panel) and of the experimental timeline (right panel). **b-d**, Left-ventricular weight normalized to tibia length (LV/TL) (**b**), isovolumic relaxation time (IVRT) (**c**) and ventricular fructose levels (**d**) of sham- or 1K1C-treated *Mlc2v-cre<sup>-</sup>* and *Mlc2v-cre<sup>+</sup>* mice injected with AAV9-fl/shSf3b1#1 viruses. **e**, Immunoblots of heart lysates from *Mlc2v-cre<sup>-</sup>* and *Mlc2v-cre<sup>+</sup>* mice operated and transduced as in **b-d** using

indicated antibodies. **f, g**, LV/TL (**f**) and ejection fraction (EF) (**g**) of sham- or TAC-treated *Mlc2v-cre<sup>-</sup>* and *Mlc2v-cre<sup>+</sup>* mice injected as indicated. **h**, Lysates of hearts from mice transduced and operated as in **f** were processed for immunoblotting with antibodies against denoted proteins. For **a-d, f, g**, number of mice per group is given in Extended Data Table 1. Error bars are s.e.m. \* $P < 0.05$ ; \*\* $P < 0.01$ ; two-tailed unpaired *t*-test (**b-d, f, g**).

Like hypertrophic growth and function, the increase in fructose levels was also dependent on Sf3b1 (Fig. 4d). Immunoblotting of lysates from these sham- or 1K1C-operated *Mlc2v-cre<sup>-</sup>* and *Mlc2v-cre<sup>+</sup>* mice confirmed that Sf3b1 was depleted where expected, preventing the generation of Khk-C isoform in ventricles of 1K1C-treated *Mlc2v-cre<sup>+</sup>* mice (Fig. 4e).

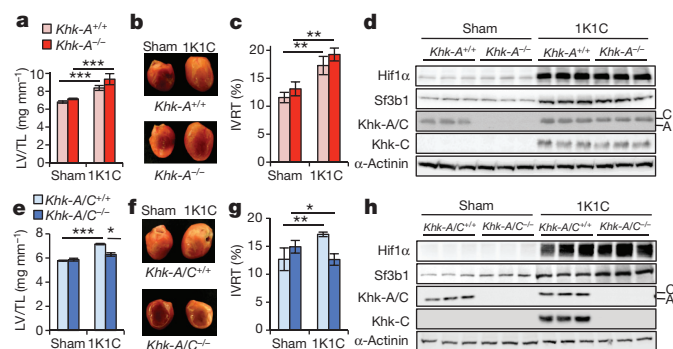
We also analysed the role of Sf3b1 in the TAC model. Application of both AAV9-fl/shSf3b1#1 and AAV9-fl/shSf3b1#2 viruses protected *Mlc2v-cre<sup>+</sup>* but not *Mlc2v-cre<sup>-</sup>* littermates from cardiac hypertrophy (Fig. 4f; Extended Data Fig. 5o and Extended Data Table 1) and systolic dysfunction (Fig. 4g; Extended Data Fig. 5p and Extended Data Table 1). Immunoblotting confirmed that Sf3b1 was downregulated and prevented the switch to the Khk-C isoform in response to TAC-induced pathologic stress in hearts of *Mlc2v-cre<sup>+</sup>* mice (Fig. 4h and Extended Data Fig. 5q). These results suggest that Sf3b1 is required for the development of pathologic cardiac hypertrophy.

### Deletion of Khk protects from cardiac growth

To assess the importance of Khk-C in cardiac hypertrophy and function in the context of pathologic stress, we investigated two genetic mouse models, which were subjected to TAC and 1K1C stresses. *Khk-A<sup>-/-</sup>* mice only express Khk-C, whereas *Khk-A/C<sup>-/-</sup>* mice are deleted for both isoforms<sup>21</sup>. Exposure of wild-type and *Khk-A<sup>-/-</sup>* mice to TAC treatment induced cardiac hypertrophy and systolic dysfunction in both mouse strains (Extended Data Fig. 6a-e and Extended Data Table 2). ATP measurements in left ventricles of sham- or TAC-treated wild-type and *Khk-A<sup>-/-</sup>* mice revealed a reduction of ATP levels in hearts after TAC compared to control mice (Extended Data Fig. 6f).

Hearts of *Khk-A/C<sup>-/-</sup>* mice, when exposed to TAC, did not display signs of hypertrophy or systolic dysfunction (Extended Data Fig. 6g-k and Extended Data Table 2). Moreover, ATP levels remained unchanged in *Khk-A/C<sup>-/-</sup>* mice exposed to TAC (Extended Data Fig. 6l). All TAC-treated hearts accumulated Hif1 $\alpha$  and Sf3b1 in the left ventricle (Extended Data Fig. 6m, n). Khk-C expression was only detectable in TAC-treated hypertrophied ventricles, but not in the corresponding hearts from *Khk-A/C<sup>-/-</sup>* mice (Extended Data Fig. 6m, n).

1K1C treatment of *Khk-A<sup>-/-</sup>* and *Khk-A<sup>+/+</sup>* mice caused significant left-ventricular growth (Fig. 5a, b; Extended Data Fig. 6o-q and Extended Data Table 3) and diastolic dysfunction (Fig. 5c and Extended Data Table 3). Hif1 $\alpha$  and Sf3b1 levels were induced in left ventricles of both 1K1C-treated strains, whereas Khk-A expression in sham-treated control mice was depleted as expected (Fig. 5d). Additional deletion of *Khk-C* protected hearts of these pathologies (Fig. 5e-g; Extended Data Fig. 6r-t and Extended Data Table 3). Immunoblotting confirmed activation of Hif1 $\alpha$  and Sf3b1 upon pathologic stress, and that Khk-C was only expressed in wild-type and *Khk-A<sup>-/-</sup>* but not in *Khk-A/C<sup>-/-</sup>* ventricles (Fig. 5h). Arterial



**Figure 5 | *Khk-A/C<sup>-/-</sup>* mice are protected from pathologic cardiac growth and contractile dysfunction.** **a-c**, Left-ventricular weight normalized to tibia length (LV/TL) (**a**), representative images of left ventricles (**b**) and isovolumic relaxation time (IVRT) (**c**) of sham- or 1K1C-operated *Khk-A<sup>+/+</sup>* and *Khk-A<sup>-/-</sup>* mice. **d**, Lysates from ventricles of sham- or 1K1C-operated *Khk-A<sup>+/+</sup>* and *Khk-A<sup>-/-</sup>* mice were processed for immunoblotting with antibodies as indicated. **e-g**, Left-ventricular weight normalized to tibia length (**e**), representative images of left ventricles (**f**) and isovolumic relaxation time (**g**) of sham- or 1K1C-operated *Khk-A/C<sup>+/+</sup>* and *Khk-A/C<sup>-/-</sup>* mice. **h**, Lysates from ventricles of sham- or 1K1C-operated *Khk-A/C<sup>+/+</sup>* and *Khk-A/C<sup>-/-</sup>* mice processed for immunoblotting with antibodies as indicated. For **a, c, e, g**, \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; two-tailed unpaired *t*-test; number of mice per group is given in Extended Data Table 2. Error bars are s.e.m.

blood pressure was similar in these mouse strains following 1K1C treatment (Extended Data Fig. 6u, v). In sum, only deletion of *Khk-A* in combination with *Khk-C*, but not *Khk-A* alone, prevents the development of pathologic growth and cardiac dysfunction upon different stresses.

## Discussion

Our results demonstrate that fructose metabolism can be provoked in the heart by pathologic stimuli through HIF1 $\alpha$  activation of SF3B1-mediated splicing of *KHK* pre-mRNA (Extended Data Fig. 7). The potential general importance of these findings draw on the central role occupied by HIF1 $\alpha$  in human disease, highlighting the HIF1 $\alpha$ -SF3B1-KHK-C axis as a promising therapeutic target.

Splicing output of the HIF1 $\alpha$  target SF3B1, which contributes to intron removal by anchoring pre-mRNA onto the spliceosome<sup>12</sup>, is sensitive to changes in the abundance of SF3B1<sup>13,22</sup>. This feature of SF3B1 to expand and/or change splice target spectrum as a function of abundance may be exploited by HIF1 $\alpha$  to affect KHK-C production. That the latter is less sensitive to changes in the levels of other factors of the U2snRNP complex investigated here implies that SF3B1 is either limiting for U2snRNP assembly/function or occupies a specific role in this context. Irrespective, this process is likely influenced by additional positive- and/or negative-acting factors, as has been demonstrated for *PKM2*<sup>23</sup>.

The criticality of KHK-C function in pathologic cardiac growth may manifest at multiple levels. Unrestrained KHK-C-driven phosphorylation of fructose may limit ATP levels and thus restrict ATP-mediated negative feedback inhibition of PFK, facilitating high glycolytic flux<sup>24</sup>. The latter may be further enhanced by an increased availability of F1P, as this metabolite has been shown to induce glucokinase and pyruvate kinase, at least in the liver<sup>25</sup>. Apart from positively affecting glucose metabolism, KHK-C and fructose availability also increased nucleic acid, lipid and protein biosynthesis, which may support growth. Although our work points to a fructose-metabolism-related function of KHK-C in cardiac hypertrophy, the source of fructose could, under certain circumstances, vary and originate alternatively from glucose<sup>15</sup>. In addition, yet-to-be-discovered functions of KHK-C unrelated to fructose metabolism (that may, for example, relate to gene regulatory activities) could contribute to cell growth control<sup>26</sup>.

Biopsies of heart disease patients demonstrate increased expression of HIF1 $\alpha$ , SF3B1 and KHK-C. Since aortic stenosis and diabetes are frequent comorbidities in senescent humans<sup>27</sup> and diabetic patients tend to have also higher fructose levels in serum<sup>28</sup>, the impact of the HIF1 $\alpha$ -SF3B1-KHK-C axis in these patients might be particularly important. With respect to other disease states, *SF3B1* is found recurrently mutated in human cancers<sup>12</sup>. It is conceivable that some of these cancer-associated mutations affect alternative splicing of *KHK* and fructose metabolism in cancer cells independent of hypoxia, thereby contributing to tumour development.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 25 June 2014; accepted 30 April 2015.**

**Published online 17 June 2015.**

1. Yip, G. W., Fung, J. W. H., Tan, Y.-T. & Sanderson, J. E. Hypertension and heart failure: a dysfunction of systole, diastole or both? *J. Hum. Hypertens.* **23**, 295–306 (2009).
2. Shiojima, I. *et al.* Disruption of coordinated cardiac hypertrophy and angiogenesis contributes to the transition to heart failure. *J. Clin. Invest.* **115**, 2108–2118 (2005).
3. Kaelin, W. G. & Ratcliffe, P. J. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Mol. Cell* **30**, 393–402 (2008).

4. Semenza, G. L. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. *Annu. Rev. Pathol.* **9**, 47–71 (2014).
5. Hill, J. A. & Olson, E. N. Cardiac plasticity. *N. Engl. J. Med.* **358**, 1370–1380 (2008).
6. Grosso, A. R. *et al.* Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res.* **36**, 4823–4832 (2008).
7. Barbosa-Morais, N. L., Carmo-Fonseca, M. & Aparício, S. Systematic genome-wide annotation of spliceosomal proteins reveals differential gene family expansion. *Genome Res.* **16**, 66–77 (2006).
8. Rappsilber, J., Ryder, U., Lamond, A. I. & Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12**, 1231–1245 (2002).
9. Diggle, C. P. *et al.* Ketohexokinase: expression and localization of the principal fructose-metabolizing enzyme. *J. Histochem. Cytochem.* **57**, 763–774 (2009).
10. Krishnan, J. *et al.* Activation of a HIF1 $\alpha$ -PPAR $\gamma$  axis underlies the integration of glycolytic and lipid anabolic pathways in pathologic cardiac hypertrophy. *Cell Metab.* **9**, 512–524 (2009).
11. Wang, C. *et al.* Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Genes Dev.* **12**, 1409–1414 (1998).
12. Bonnal, S., Vigevari, L. & Valcárcel, J. The spliceosome as a target of novel antitumour drugs. *Nature Rev. Drug Discov.* **11**, 847–859 (2012).
13. Corriero, A., Minana, B. & Valcárcel, J. Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. *Genes Dev.* **25**, 445–459 (2011).
14. Dolatshad, H. *et al.* Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells. *Leukemia* **29**, 1092–1103 (2015).
15. Ramasamy, R. & Goldberg, I. J. Aldose reductase and cardiovascular diseases, creating human-like diabetic complications in an experimental model. *Circ. Res.* **106**, 1449–1458 (2010).
16. Lewis, C. A. *et al.* Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. *Mol. Cell* **55**, 253–263 (2014).
17. Cirillo, P. *et al.* Ketohexokinase-dependent metabolism of fructose induces proinflammatory mediators in proximal tubular cells. *J. Am. Soc. Nephrol.* **20**, 545–553 (2009).
18. Fang, M. *et al.* The ER UDPase ENTPD5 promotes protein N-glycosylation, the Warburg effect, and proliferation in the PTEN pathway. *Cell* **143**, 711–724 (2010).
19. Vander Heiden, M. G. *et al.* Evidence for an alternative glycolytic pathway in rapidly proliferating cells. *Science* **329**, 1492–1499 (2010).
20. Sharma, S. *et al.* Intramyocardial lipid accumulation in the failing human heart resembles the lipotoxic rat heart. *FASEB J.* **18**, 1692–1700 (2004).
21. Diggle, C. P. *et al.* Both isoforms of ketohexokinase are dispensable for normal growth and development. *Physiol. Genomics* **42A**, 235–243 (2010).
22. Gao, Y., Trivedi, S., Ferris, R. L. & Koide, K. Regulation of HPV16 E6 and MCL1 by SF3B1 inhibitor in head and neck cancer cells. *Sci. Rep.* **4**, 6098 (2014).
23. David, C. J., Chen, M., Assanah, M., Canoli, P. & Manly, J. L. HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature* **463**, 364–368 (2010).
24. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 (2009).
25. Shiota, M., Galassetti, P., Monahan, M., Neal, D. W. & Cherrington, A. D. Small amounts of fructose markedly augment net hepatic glucose uptake in the conscious dog. *Diabetes* **47**, 867–873 (1998).
26. Luo, W. *et al.* Pyruvate Kinase M2 Is a PHD3-Stimulated Coactivator for Hypoxia-Inducible Factor 1. *Cell* **145**, 732–744 (2011).
27. Falcão-Pires, I. *et al.* Diabetes mellitus worsens diastolic left ventricular dysfunction in aortic stenosis through altered myocardial structure and cardiomyocyte stiffness. *Circulation* **124**, 1151–1159 (2011).
28. Kawasaki, T., Akanuma, H. & Yamanouchi, T. Increased fructose concentrations in blood and urine in patients with diabetes. *Diabetes Care* **25**, 353–357 (2002).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank S. Georgiev, T. Simka, S. Xu, C. Bischoff, J. M. Dominguez, W. Kovacs and M. Piontek and other members of the Krek laboratory for discussions, help and technical assistance. We are grateful to M. Stoffel for performing tail vein injections. K. Chien, A. Asipu and R. J. Johnson provided mouse lines. This work was supported by grants from Sinergia (Swiss National Science Foundation) to W.K., T.P. and J. U. and the Swiss Heart Foundation to W.K.

**Author Contributions** P.M., J.K. and W.K. designed and P.M. executed most experiments. F.G. performed ketohexokinase assays, *ex vivo* glucose/fructose uptake and lipid loading studies. A.S. performed mouse surgeries, echocardiography and necropsy analysis under the supervision of T.P., M.H. and N.Z. performed metabolomic analysis. J.U. and M.K. performed splice junction microarrays and initial data analysis. C.C. performed ATP measurements and Sf3b1-rescue experiments, O.F. quantified lipids and A.V. performed biodistribution experiments. N.F. and Y.C. generated the splice factor list and analysed splice junction microarray data. S.S. and S.N.S. provided human left ventricular biopsies. P.M. and W.K., with help from J. K., wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.K. ([wilhelm.krek@biol.ethz.ch](mailto:wilhelm.krek@biol.ethz.ch)).



## METHODS

**Human ventricular biopsies.** Left-ventricular tissues were obtained from patients with hypertrophic cardiomyopathy and aortic stenosis (Supplementary Table 4) and healthy controls. Additional left-ventricular biopsies were provided by W. Linke (University of Münster). Clinical and demographical data from the corresponding patients were previously published<sup>29</sup>. All procedures were in compliance with local ethical committees. An informed consent was obtained from all patients included in the study.

**Mouse strains.** *Khk-A*<sup>-/-</sup> and *Khk-A/C*<sup>-/-</sup> mice were obtained from A. Asipu (University of Leeds) and R. Johnson (University of Colorado). *MLC2v-cre* line was from J. Chen (University of California). All mice were maintained in a conventional animal facility at RCHCI and EPIC, ETH Zürich. Maintenance and animal experimentation were in accordance with the Swiss Federal Veterinary Office (BVET) guidelines. Data presented in this manuscript represent studies with male mice. After baseline echocardiography mice were randomly distributed to sham or treatment groups. Animal surgeries and echocardiography was performed blinded, that is the involved veterinarian was not aware of the genotype of the mice before or during operation. Animal numbers for experiments were chosen based on expected mortality rates and phenotypical or functional changes of hearts in wild-type mice due to treatment/surgery. Animals were excluded from the study in case of death before the experimental endpoint or based on the evaluation of pain using a standardized score sheet, approved by BVET.

**One-kidney, one-clip.** Six- to seven-week-old *Khk-A*<sup>-/-</sup>, *Khk-A/C*<sup>-/-</sup> and corresponding *Khk-A*<sup>+/+</sup> and *Khk-A/C*<sup>+/+</sup> male mice with a pure C57/BL6J background as well as *Mlc2v-cre*<sup>+</sup> and *Mlc2v-cre*<sup>-</sup> male mice with a C57/BL6J 129S mixed background were anaesthetized and surgery was performed as previously described<sup>10</sup> by inhalation of isoflurane (4% in oxygen, 1 l min<sup>-1</sup>). A sham procedure, which included the entire surgery with the exception of artery clipping, was applied to control mice.

**Trans-aortic constriction.** 10–12-week-old *Khk-A*<sup>-/-</sup>, *Khk-A/C*<sup>-/-</sup> and corresponding *Khk-A*<sup>+/+</sup> and *Khk-A/C*<sup>+/+</sup> male mice as well as *Mlc2v-cre*<sup>+</sup> and *Mlc2v-cre*<sup>-</sup> male mice were subjected to pressure overload by trans-aortic constriction (TAC) through constriction of the descending aorta as described<sup>30</sup>. The mice were monitored for 14–18 days after surgery, and their heart functions determined by echocardiography.

**In vivo isoproterenol delivery.** Alzet model 2002 mini-osmotic pumps (Alzet) were filled with isoproterenol, and were set to deliver at 30 mg kg<sup>-1</sup> per day for 14 days in C57/BL6J wild-type mice. Pumps were then implanted into 9–10-week-old male mice under isoflurane anaesthesia as previously described<sup>10</sup>.

**Echocardiography.** Transthoracic echocardiography was performed using a 30 MHz probe and the Vevo 770 Ultrasound machine (Visualsonics) under mild anaesthesia with 1.5% isoflurane as previously described<sup>31</sup>.

**Direct arterial blood pressure measurements in mice.** The arterial blood pressure measurements in mice were performed as described previously<sup>32</sup>. Blood pressure was monitored for about 1 h in awake animals that were unconstrained.

**Isolation and maintenance of primary neonatal mouse cardiomyocytes.** Isolation of neonatal mouse cardiomyocytes was performed as previously described<sup>10</sup>. Experiments involving HIF1 $\alpha$ , SF3B1 and KHK-C were performed with maintenance medium supplemented with physiologic concentrations of glucose (5 mM) and fructose (25  $\mu$ M). The medium was changed daily. Pharmacological treatment was done with phenylephrine at a concentration of 100  $\mu$ M and with isoproterenol at 10  $\mu$ M for 48 h before preparation of SDS samples or processing for RNA isolation.

**Lentivirus production and transduction.** Lentiviral particles were produced in HEK-293T cells as previously described<sup>33</sup>. HEK-293T cells were purchased from American Type Culture Collection (ATCC) and regularly checked for mycoplasma contamination using a PCR-based mycoplasma detection kit (Sciencell). Cells were transduced with lentivirus 3 h after isolation by adding virus-containing supernatants to the cell culture medium and incubating cells at 37 °C/5% CO<sub>2</sub> overnight.

**Transient transfection.** The pcDNA3 SF3b1 construct was transiently transfected by using the Lipofectamin 2000 (Invitrogen) in 2-day-old NMCs according to the manufacturer's protocol. Cells were harvested and analysed 36 h after transfection.

**AAV9-fl/fl-shSF3b1.** For the generation of the viral constructs, we cloned two shRNAs targeting *Sf3b1* in a pSico vector where the U6 promoter driven shRNA expression is Cre-lox regulated<sup>34</sup>. The shRNA sequences were the following: AAV9-fl/fl-shSF3b1#1, sense (5'–3'): TGCATCGTCTCCAAAGATTTCAGAGAGAAATCTTTGGAGGACGATGCTTTTTC, antisense (5'–3'): TCGAGAAAAAGCATCGTCTCCAAAGATTTC TT GAAATCTTTGGAGGACGATGCA; AAV9-fl/fl-shSF3b1#2, sense (5'–3'): TG CTGCCTACGTCAGATTATCAAGAGATAATCTGACGTAAGGCAGCTTTTTC, antisense

(5'–3'): TCGAGAAAAAGCTGCCTACGTCAGATTATCTCTTGAATAA TCTGACGTAAAGGCAGCA.

After *in vitro* validation of shRNA efficiency using AdenoCre expression to induce Cre-mediated recombination and activation of shSF3b1 expression, the region bearing the U6 promoter, the shRNA and the TATA-lox sites of the pSico construct was amplified and ligated into NheI/XhoI linearized AAV-bGH(+) vector. Viral amplification was done by Vectorbiolabs. AAV9 viruses were injected at a concentration of  $1.2 \times 10^{13}$  genome copies (GC) per kg body weight into tail veins of *Mlc2v-cre*<sup>+</sup> or *Mlc2v-cre*<sup>-</sup> mice.

**Plasmid constructions.** The pLenti-pgk-HIF1 $\alpha$ ΔODD puro lentiviral expression vector was generated by subcloning the HIF1 $\alpha$ ΔODD fragment from a pcDNA3-HA-HIF1 $\alpha$ ΔODD(401–603) plasmid<sup>35</sup> as described previously<sup>36</sup>. The KHK-A and KHK-C overexpression constructs were generated by subcloning KHK-A and KHK-C from the pET11a-KHK-A and pET11a-KHK-C expression constructs (provided by A. Asipu) into the Gateway Entry vector pENTR1a and by LR-reaction in the pLenti-pgk-puro lentiviral expression vector. As a corresponding control the pLenti-pgk-puro empty vector was used. SF3A3 and SF3B3 overexpression constructs were generated by subcloning SF3A3 (from human ORFeome 8.1 collection, Thermo Scientific) or SF3B3 (from Addgene, plasmid #11241) in the pLKO.1 CMV puro construct provided by A. Ittner (ETH Zürich), by using EcoRI and SalI (SF3A3) or EcoRI and MluI (SF3B3) restriction sites. As a control vector the empty pLKO.1 CMV puro construct was used. Since the wild-type coding sequence of SF3b1 is toxic to bacteria, we used a codon optimized SF3b1 to perform overexpression experiments (the DNA and amino acid sequence can be found in Supplementary Data 1). First, we amplified SF3b1 from mouse cDNA until the internal XhoI restriction site (1–1,568 bp) and subcloned it into a pcDNA3 vector by using EcoRI/XhoI restriction sites. The synthesis and codon optimisation of the second part (1,578–3,915 bp) was performed by Genscript. Finally we subcloned the two parts of SF3b1 by double ligation in a pcDNA3 vector by using EcoRI, XhoI and XbaI restriction sites.

**Sf3b1 promoter-luciferase construct.** 1.5 kilobases (kb) of the *Sf3b1* promoter was amplified from mouse BAC genomic DNA and cloned into the pGL3 luciferase reporter vector (Stratagene). HRE mutants of the respective promoters were generated by recombinant PCR<sup>37,38</sup>. Sense and antisense primers were generated bearing HRE mutations, which were used to amplify the mutant 5' and 3' regions of the promoter, respectively. The 5' and 3' products generated from the respective PCR reactions were pooled afterwards at a 1:1 ratio and re-amplified using primers targeting the 5' and 3' ends of the respective promoters. To check for the correctness of the HRE mutation in the right direction, HRE mutants of the respective promoters were verified by sequencing (Microsynth).

**shRNA knockdown of Khk-A and Khk-C.** Khk-A- or Khk-C-specific shRNAs were designed using the Life Technologies BLOCK-iT RNAi Designer online tool (Life Technologies), searching for 21mer targets of either *Khk-A*- or *Khk-C*-specific transcripts. shRNAs were automatically compared to the mouse database using BLAST. As a loop sequence, the TRC-library loop sequence TCTGAG was used. Per isoform, three different shRNAs were chosen as promising, according to the program's algorithm and cloned into the pLKO.1 lentiviral vector. The presence of the shRNA was confirmed by sequencing (Microsynth). shRNAs were then tested in NMCs for knockdown efficiency and specificity for either Khk-A or Khk-C. After validation, shRNAs shKhk-A#3, referred as shKhk-A and shKhk-C#2, referred as shKhk-C were chosen for all experiments described. The sequences are as follows: shKhk-A, sense (5'–3'): CCGGGATGACCTCCGCC AACAT TCTCTCGAGAGAATGTTGGCGGAGGTACATCTTTTGTG; shKhk-C, sense (5'–3'): CCGGGCTCTTGCTGCATCGTCAACACTCGAGTGTGAC GATGCAGCAAGAGC TTTTGTG.

**TRC lentiviral shRNA constructs.** For lentiviral shRNA mediated knockdown the following shRNAs from the TRC library were used in a pLKO.1 construct: Sf3b1 (shSF3b1#1, TRCN0000295293; shSF3b1#2, TRCN0000295294), Sf3a3 (shSF3a3#1, TRCN0000285834; shSF3a3#2, TRCN0000277028), Sf3b2 (shSF3b2#3, TRCN0000123811; shSF3b2#4, TRCN0000123812), Sf3b3 (shSF3b3#5, TRCN0000311577; shSF3b3#11 TRCN0000109223), Glut5 (shGlut5#4, TRCN0000079392; shGlut5#5, TRCN0000079388), Hif1 $\alpha$  (shHif1 $\alpha$ , TRCN0000232220), Khk (shKhk, TRCN0000416511). As a non-targeting shRNA construct the pLKO.1 vector SHC002 (Sigma) was used.

**RNA-isolation and reverse transcription.** RNA was isolated with the RNeasy Mini Kit (Qiagen) and reverse-transcribed using a EcoDry Premix (OligoDT) (Clontech) according to the manufacturer's instructions. Quantitative real-time PCR (qRT-PCR) reactions were prepared with Roche LightCycler 480 SYBR Green I Master as recommended by the manufacturer and run on a LightCycler 480 (Roche). C<sub>t</sub> values were normalized to the housekeeping gene *Hprt1*. Primer sequence of qRT-PCR primers used in this study are as follows. Mouse: *Hprt1* 5'-GGGGCTGTACTGCTTAACCAG-3' and 5'-TGC TCATCAGTTGCCACTTC-3'; *Sf3b1* (primer pair A in Extended Data Fig. 3a)

5'-GCATCGTCCTCCAAAGATTGC-3' and 5'-GCATTCATTTTAGGATCG GGAGT-3'; Sfb1 (primer pair B in Extended Data Fig. 3a) 5'-CTCATGG GCTGTGCTATCTTGC-3' and 5'-GTTGCTGCCTCAGCCAAAGC-3'; Sfb1 (codon optimized, primer pair C in Extended Data Fig. 3a) 5'-CGTTAAA GGCCATCGTGAACG-3' and 5'-CAATTTTCCTGCACTTTCTCATGC-3'; Sfb3 5'-GAACGGCTTATGGATGTTATGGC-3' and 5'-GGTGATCGGAA TTGATCTGGTC-3'; Sfb2 5'-CATCCGAACCTCCCAAAGG-3' and 5'-GC TCCT ATCTCTGCCAATCTGG-3'; Sfb3 5'-ATTTGGAGTTACTACGCC AG-3' and 5'-GCCATAAGTGACCGGATAACAC-3'; Khk-A 5'-CCAACATT CTGTGGAACCTACG-3' and 5'-CCTTCTCAAAGTCTTAGCAG-3'; Khk-C 5'-GCTGACTTCAGGCAGAGG-3' and 5'-CCTTCTCAAAGTCC TTAGC AG-3'; Glut-1 5'-CAGTTCGGCTATAACACTGGTG-3' and 5'-GCCCCGAC AGAGAAGAT-3'; Pdk1 5'-GGACTTCGGGTCAAGTGAATGC-3' and 5'-CGC AGAAACATAAACGAGGTC-3'; Pkm2 5'-CGCCTGGACATTGACTCTG-3' and 5'-GAAATTCAGCCGAGCCACATT-3'; Glut5 5'-CCAATATGGGTACA ACGTAGCTG-3' and 5'-CCAATATGGGTACAACGTAGCTG-3'; sk- $\alpha$ -Actin 5'-TATGTGGCTATCCAGGCGGTG-3' and 5'-CCCAGAATCCAACAGCAT GC-3'; Nppa 5'-GTGCGGTGTCCAA CACAGAT-3' and 5'-TCCAATCTGTCT AATCCTACCC-3'; Nppb 5'-AGTCCTTCG GTCTCAAGGCA-3' and 5'-CCG ATCCGCTCTATCTGTGC-3';  $\alpha$ Mhc 5'-GCCCAGTACCTCCGAAAGTC-3' and 5'-GCCTTAACATACTCCTTGTG-3';  $\beta$ Mhc 5'-GCATTCTCCTGCTGTT TCCTT-3' and 5'-TGGATTCTCAAACGTGTCTAGTGA-3'; human: HPRT1 5'-CTGGCGTCGTGATTAGTATG-3' and 5'-AGACGTTCAAGTCTGTCC ATAA-3'; SF3B1 5'-GTGGGCTCGATTCTACAGG-3' and 5'-GATGTCAC GTATCCAGCAAATCT-3'; KHK-A 5'-GCTATTCTGTGGACCTACGCT-3' and 5'-CAATGTGGATCCACTTGAAGT-3' and KHK-C 5'-CTCCTGCT GCATCAAC-3' and 5'-CAATGTGGATCCACTTGAAGT-3'; GLUT1 5'-GGCCAAGAGTGTGCTAAAGAA-3' and 5'-ACAGCGTTGATGCCAGA CAG-3'; PDK1 5'-CTGTGATACGGATCAGAAACCG-3' and 5'-TCCACCAA ACAATAAAGAGTGCT-3'; GLUT5 5'-TCCAGTATGGGTACAACGTGG-3' and 5'-CATGGACACGGTTACAGACCA-3'; AKR1B1 5'-TTTTCCCATTTGG ATGATCGG-3' and 5'-CCTGGAGATGGTTGAAGTTGG-3'; NPPA 5'-CA ACGCAGACCTGATGGATT-3' and 5'-AGCCCCGCTTCTTCTATTC-3'; NPPB 5'-TGGAAACGTCCGGGTTACAG-3' and 5'-CTGATCCGGTCCAT CTTCCT-3';  $\alpha$ MHC 5'-GCCCTTTGACATTGCACTG-3' and 5'-CGGGA CAAAATCTTGCTTTGA-3';  $\beta$ MHC 5'-ACTGCCGAGACCGAGTATG-3' and 5'-GCGATCCTTGAGTTGTAGAGC-3'.

**Splice-junction microarrays.** Total RNA was isolated from NMCs infected with shNs and shHif1 $\alpha$  lentiviruses and cultured in either 21% or 3% O<sub>2</sub>. In another experiment NMCs were transduced with shNs and shSF3b1 lentiviruses and cultured at 3% O<sub>2</sub> before RNA was isolated. Complementary DNA samples were prepared using the GeneChip WT cDNA Synthesis and Amplification Kit (Affymetrix). cDNA was hybridized to AltSplice splice-junction microarrays and resulting data were analysed with version 3 of ASPIRE software as previously described<sup>39</sup>. The unprocessed cel files with microarray data are available from ArrayExpress with accession numbers E-MTAB-2652 and E-MTAB-3203, respectively.

**Luciferase assays.** The luciferase assay was performed from material of NMCs using a luciferase kit (Promega) as recommended by the manufacturer and analysed on the MicroLumatPlus LB 96V (Berthold Technologies).

**Chromatin and RNA immunoprecipitations.** ChIP assays were performed with material from NMCs and the assays carried out using the ChIP-IT Kit (Active Motif) according to the manufacturer's instructions. ChIP was performed with a ChIP-grade antibody from Abcam. *In silico* promoter analyses were performed using MatInspector and DiAlignTF (Genomatix). Sequences of primers used are 5'-TTTCGCTGGCTTTTAAAGGA-3' and 5'-ATCTCGCCATTTGTCCAC-3'.

RNA immunoprecipitation assays were performed with material from NMCs using a RNA-ChIP-IT kit from Active Motif as recommended by the manufacturer. RNA immunoprecipitation was performed using a Sf3b1 antibody from MBL International. Sequences of primers used to detect potential binding of Sf3b1 on the branch points of *Khk-A* and *Khk-C* are the following: branch point primer for *Khk-A*, 5'-CTACTTCCCTCTGCTCAGCC-3' and 5'-GGTCATCC AGACAAAATGC-3'; branch point primer for *Khk-C*, 5'-GGACCCCTTCTCTGAGGC-3' and 5'-CACTTGAGACACATCCACGC-3'.

**Immunoblotting.** Dissected hearts and NMCs lysates were homogenized by freeze slamming and solubilized in a modified SDS sample buffer sonicated and boiled for 5 min<sup>40</sup>. Protein lysates were resolved on 8%, 10% or 15% polyacrylamide minigels (BioRad) and transferred onto nitrocellulose membrane (GE Healthcare) by wet or semi-dry transfer. The following antibodies were used for immunoblotting rabbit HIF1 $\alpha$  (catalogue no. NB-100-47; Novus Biologicals), a rabbit SF3B1 (catalogue no. ab172634; Abcam), a rabbit-KHK (catalogue no. HPA007040; Sigma), a rabbit KHK-C (custom made by Eurogentec) a rabbit Cre (catalogue no. 69050; Novagen), a rabbit SF3A3 (catalogue no. ab140605),

a mouse SF3B2 (catalogue no. ab56800), a rabbit SF3B3 (catalogue no. NBP2-20258) mouse  $\alpha$ -actinin (catalogue no. A7811; Sigma) and a chicken-GFP (catalogue no. GFP-1020; Aves Labs). Immunodetection and visualization of signals by chemiluminescence was carried out as described<sup>40</sup>, or by using the ImageQuant LAS 4000 mini (GE Healthcare). Immunoblots for Khk-A/C and Khk-C were visualized using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

**Ketohexokinase splicing assay I.** This *Khk* splicing assay including amplification of part of the *Khk* cDNA and the identified alternative spliced region was adapted from a splice assay for *PKM1/M2* isoforms<sup>23</sup>. RNA isolated from NMCs transduced with shNs or shSf3b1 lentiviruses after treatment with isoproterenol or HIF1 $\alpha$ ODD and RNA isolated from sham- or 1K1C-treated wild-type left ventricles was reverse-transcribed then amplified using a forward primer (5'-GAAGCAGATCCTGTGCGTG-3') that bound to shared *Khk* A/C exon 1 and a reverse primer (5'-ATTGTGCTCCTCTATCCGC-3') that bound to shared *Khk* A/C exon 5 sequences. The 468-bp PCR products representing both *Khk-A* and *Khk-C* transcripts were then digested with HincII and run on a 1.5% agarose gel, after which the 468 bp (uncut *Khk-A*-specific amplicon), 294 bp and 174 bp (HincII-cleaved *Khk-C*-specific amplicon) products were quantitated using ImageJ software.

**Ketohexokinase splicing assay II.** To analyse *KHK* isoforms in human and mouse hearts or NMCs RT-PCR was performed. Total RNA was isolated from hearts or NMCs and reverse-transcribed into cDNA using the RNA to cDNA EcoDry Premix (OligoDT) kit (Clontech). Full length *KHK-A/C* cDNA was amplified by RT-PCR, loaded onto an agarose gel and purified by gel extraction. One hundred nanograms of full-length *KHK-A/C* cDNA was used as a template to amplify *KHK-A* and *KHK-C* specific fragments, respectively, using primer pairs consisting of a common *KHK* A/C forward and *KHK-A*- or *KHK-C*-specific reverse primer set. To be able to distinguish between *KHK-A* and *KHK-C* isoforms, the corresponding amplicons differed in length. The influence of fragment length on fragment amplification was minimised by designing two sets of primer pairs for each isoform (primer set 1 vs primer set 2): *Khk-A*<sub>1</sub> < *Khk-C*<sub>1</sub> and *Khk-A*<sub>2</sub> > *Khk-C*<sub>2</sub>. The primer sequences are the following: primer pair 1 (human): *KHK-A/C* 5'-GAGAAGCAGATCCTGTGC-3', *KHK-A* 5'-AGCGTAGGTCC ACAGAATAGC-3' and *KHK-C* 5'-GGAGTTGTTGATGATGCAGC-3'; primer pair 2 (human): *KHK-A* 5'-GCTATTCTGTGGACCTACGCT-3' and *KHK-C* 5'-CAGTTCAGTGGATCCACATTG-3'; *KHK-A/C* 5'-GGTGTGTGTGCGTCTATCC-3'; primer pair 1 (mouse): *Khk-A/C* 5'-GAAGCAGATC CTGTGCGTG-3', *Khk-A* 5'-GGTCTGAAGGACCACATATCG-3' and *Khk-C* 5'-CATTGGAGTTGTTGACGATGC-3'; primer pair 2 (mouse): *Khk-A* 5'-CAT TCTGTGGAAGTACGATATGTG-3' and *Khk-C* 5'-GTACCATTATACTCT ACGACACGAAC-3'; *Khk-A/C* 5'-GGTGTGTGTGCGTCTATCC-3'.

**Ketohexokinase activity assay.** The ketohexokinase activity assay<sup>41</sup> was modified by adding Roche Complete Protease Inhibitor Cocktail and Roche PhosSTOP to the reaction mixture, addition of 140  $\mu$ g NMCs protein per reaction and incubating for 2 h at 37 °C. Spectrophotometric absorbance measurements of the remaining fructose were performed in 96-well plates at 515 nm using a Spectra MAX 190 plate reader (Molecular Devices). Wild-type liver protein was used as a positive control and reactions without protein were used as negative controls.

**Metabolomics.** 4 × 10<sup>5</sup> NMCs were cultured per 3-cm dish. Cells were washed with PBS and whole-cell culture plates were snap frozen on dry ice. Metabolites were extracted twice from plates with hot (75 °C) 70% ethanol. Targeted analysis of hexose-phosphates levels and of their <sup>13</sup>C-labelling pattern was performed by ion-pairing, reverse phase, ultra-performance liquid chromatography-tandem mass spectrometry on a Thermo Quantum Ultra instrument (Thermo Scientific)<sup>42</sup>. Technical quadruplicates were merged using their mean.

Untargeted analysis of metabolites was performed by flow injection-time-of-flight mass spectrometry on a Agilent 6550 Instrument in negative mode 4 GHz, high resolution in a *m/z* range of 50–1,000 as reported earlier<sup>43</sup>. Ions were annotated to ionic derivatives of the metabolites listed in the KEGG database (organism *hsa*) using 0.001 atomic mass units (a.m.u.) tolerance. Data processing and analysis was done with Matlab.

**Immunofluorescent stainings.** Immunofluorescent stainings were performed as described previously<sup>10</sup>. Cells were permeabilized and incubated with sarcomeric  $\alpha$ -actinin antibody (Sigma) diluted 1:500 in 2% horse serum, 0.025% TritonX-100/PBS overnight at 4 °C. Cells were washed with PBS and incubated with 4',6-diamidino-2-phenylindole (DAPI), Phalloidin and AlexaFluor488 anti-mouse secondary antibody for 1 h at room temperature. Afterwards, stainings were washed and mounted in Lisbeth medium.

**[<sup>3</sup>H]leucine incorporation assay.** Cardiomyocyte hypertrophy was evaluated by measuring [<sup>3</sup>H]leucine incorporation into newly synthesized proteins<sup>44</sup>. Cells were counted in a Neubauer chamber after trypsinization and scintillation counts were normalized to absolute cell number.

**Ex vivo [ $^3\text{H}$ ]fructose/[ $^{14}\text{C}$ ](deoxy)-glucose uptake.** Animals were euthanized 8 weeks after surgery and hearts were shortly placed into 1 M KCl solution. The left ventricle was cut into equal tissue pieces, dried on tissue, weighed and placed into 6-well plates with 2 ml medium (serum-free DMEM, 5 mM glucose and 8  $\mu\text{M}$  fructose). Tissue explants were incubated in the same medium containing 1 mCi ml $^{-1}$  [ $^3\text{H}$ ]fructose (specific activity 5 Ci mmol $^{-1}$ ) and 0.1 mCi ml $^{-1}$  [ $^{14}\text{C}$ ]deoxy-D-glucose (specific activity 58 mCi mmol $^{-1}$ ) and harvested at respective time points. Harvested explants were washed two times with PBS and dissolved in 0.5 ml 1 M KOH (70 °C, 10 min, shaking). 5 ml scintillation liquid was added and radioactivity was measured in a Beckman LS6500 scintillation counter. The radioactivity counts were normalized to tissue weight.

**In vivo [ $^3\text{H}$ ]fructose/[ $^{14}\text{C}$ ](deoxy)-glucose uptake and biodistribution.** A mixture of 20  $\mu\text{l}$  1 mCi ml $^{-1}$  [ $^3\text{H}$ ]fructose (specific activity 5 Ci mmol $^{-1}$ ) and 200  $\mu\text{l}$  0.1 mCi ml $^{-1}$  [ $^{14}\text{C}$ ]deoxy-D-glucose (specific activity 55 mCi mmol $^{-1}$ ) was applied to sham- or TAC-treated C57/Bl6 J mice at 3 weeks post operation by oral gavage. After 45 min animals were weighed and euthanized. Denoted organs were collected and weighed. Tissue was digested with 4 M ethanolic KOH, radioactivity was measured in a Beckman LS6500 scintillation counter (Beckman) and calculated as disintegrations per minute (d.p.m.) per mg tissue weight.

**Fructose content.** Heart fructose content was measured using the fructose colorimetric/fluorometric assays kit according manufacturer's instructions (Biovision) and as described previously<sup>45</sup>. Values were normalized to protein content.

**[ $^3\text{H}$ ]fructose incorporation into RNA, DNA and protein.**  $4 \times 10^5$  NMCs were seeded in 3-cm dishes and transduced in quadruplicates. On day 3, NMCs were incubated in maintenance medium containing 5 mM glucose, 8  $\mu\text{M}$  unlabelled Fructose and 1 mCi ml $^{-1}$  [ $^3\text{H}$ ]fructose (specific activity 5 Ci mmol $^{-1}$ ) overnight. NMCs were washed three times with PBS before cell lysis was performed with 0.7 ml TRIzol (Invitrogen). Fractionation into RNA, DNA and protein was performed according to the manufacturer's instructions with adaptations of the volumes according to the starting volume.

**ADP/ATP Assay.**  $5 \times 10^4$  NMCs were cultured in a 96-well format and treated as denoted. After 3 days in culture NMCs were analysed using a Bioassay Systems EnzyLight ATP:ADP ratio kit (ELDT-100, Bioassay Systems) according to manufacturer's instructions. Cellular ATP from 10  $\mu\text{l}$  of lysate reacted with D-luciferin to generate a light signal that was proportional to the intracellular ATP concentration. In a second step, ADP was converted to ATP by an enzymatic reaction, and reacting with D-luciferin as in the first step. Luminometry was analysed using a MicroLumatPlus LB 96V (Berthold Technologies).

**Myocardial ATP content.** The level of myocardial ATP was measured by using a colorimetric/fluorometric assay kit (Abcam) according to the manufacturer's instructions. Left-ventricular samples of about 10 mg weight were used as a starting material. The absorbance was read at 550 nm using a Tecan Infinite M1000 Pro plate reader and ATP content was calculated based on a standard curve generated from ATP-standards on the same 96-well plate. Values were normalized to protein concentrations obtained from Bradford assays.

**[ $^3\text{H}$ ]fructose incorporation into lipids.**  $4 \times 10^5$  cells were seeded in 3-cm dishes and transduced in quadruplicates. On day 3, cells were loaded with 0.25 mM oleic acid overnight and 1  $\mu\text{Ci}$  ml $^{-1}$  [ $^3\text{H}$ ]fructose (specific activity 5.0 Ci mmol $^{-1}$ ) was added to the medium. On day 4, cells were washed twice with PBS and lipids were extracted by hexane-isopropanol extraction<sup>46</sup>. Protein concentration was measured by BCA protein assay and radioactivity count was normalized to protein content.

**Oil Red O staining.** Oil Red O staining was combined with immunofluorescence according to a protocol in ref. 47. Spot-detection was performed with the Anchor image analysis platform (<http://www.microscopyimages.org/>). Nuclei are segmented by first thresholding the (Gaussian-blurred) DAPI channel with Otsu's method, and then a watershed segmentation on the inverted Euclidian distance transform. Neighbouring segments are merged to increase an ellipticity score. A merge occurs if:  $\text{ell}(xy) > 0.5[\text{ell}(x) + \text{ell}(y)]$ , where  $xy$  is the merge of segments  $x$  and  $y$ . Segments with an area  $> 50 \mu\text{m}^2$  are deemed nuclei. Cells are segmented by thresholding the histogram-stretched FITC channel, combining with the nuclei binary-image, and performing ultimate erosion using the segmented nuclei as seeds. Spots are segmented by first performing rolling-ball background correction, thresholding, and then a watershed segmentation on the inverted intensity after median filtering. Segments with an area  $> 0.2 \mu\text{m}^2$  and a minimum gradient signal are deemed spots.

**Nile red staining.**  $4 \times 10^5$  NMCs were incubated with 0.25 mM oleic acid (Sigma-Aldrich) overnight and stained with 10  $\mu\text{g}$  ml $^{-1}$  Nile red (Sigma-Aldrich) for 10 min in PBS<sup>48</sup>. Values were normalized to cell number.

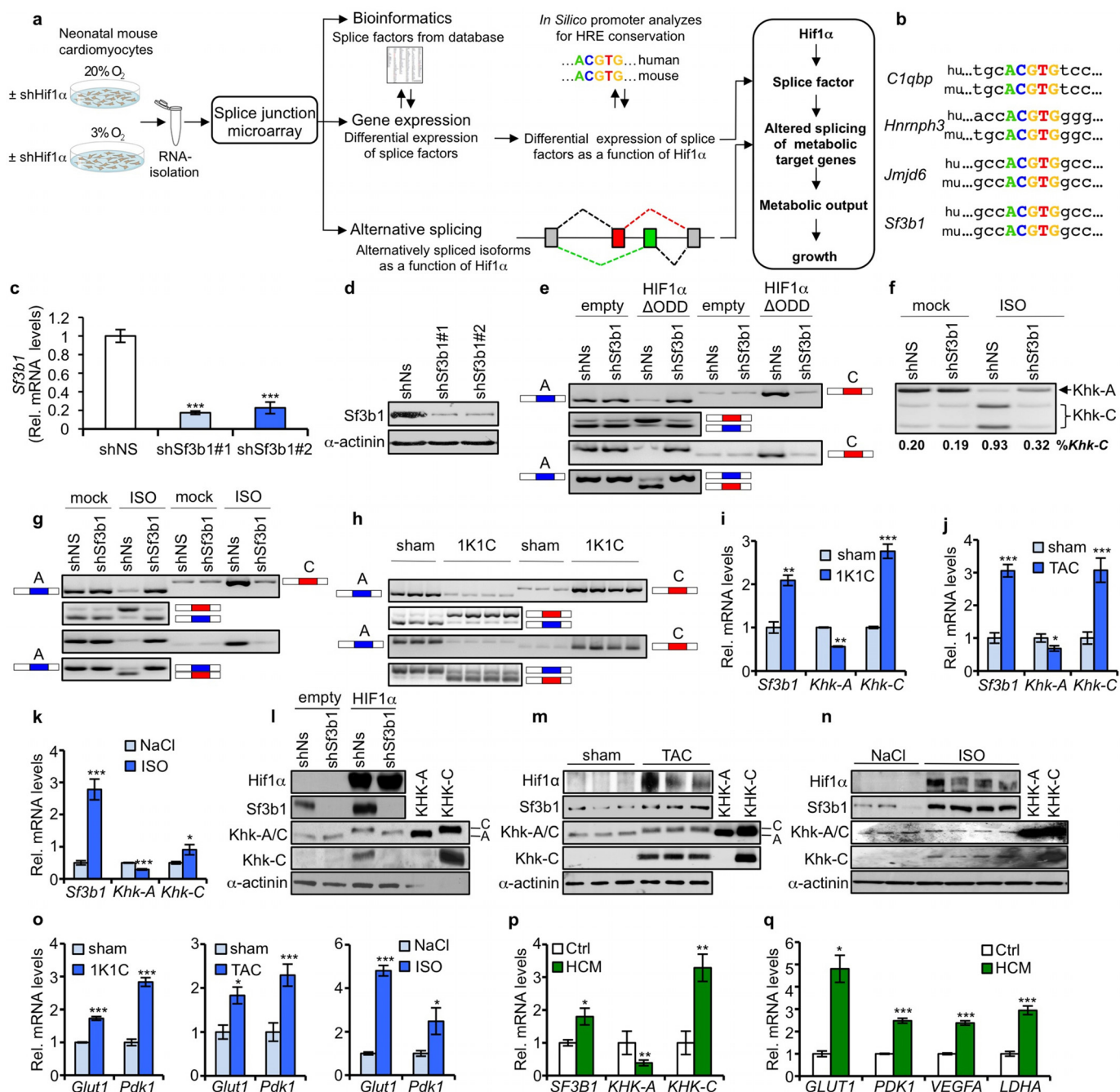
**Cell size quantification.** NMCs were stained after 3 days of culturing for sarcomeric  $\alpha$ -actinin (1:500, Sigma), phalloidin and DAPI. Pictures of all channels were taken using a 20 $\times$  magnification. Using the software Cell Profiler, sarcomeric  $\alpha$ -actinin-positive cells were identified and cell area was quantified. Multi-nucleated cells were counted manually and average area per cell was corrected taking into account the number of multi-nucleated cells.

**Measurements of extracellular acidification and oxygen consumption rate.**  $1 \times 10^5$  NMCs per well were seeded in a 24-well Seahorse cell culture plate (SeaHorse Bioscience) and assayed at day 3. ECAR and OCR were measured in parallel using a SeaHorse XF24 Flux Analyzer (SeaHorse Bioscience) in NMCs incubated in low-buffered Krebs-Henseleit buffer, supplemented with glucose (5 mM) and fructose (25  $\mu\text{M}$ ). After five consecutive basal measurements of ECAR to measure glycolysis related lactate production, 0.25 mM oleic acid was added and oxygen consumption rates were measured again for five consecutive times. Finally, 1  $\mu\text{M}$  carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone was injected to stimulate maximal OCR in NMCs, reflecting mitochondrial reserve capacity.

**Statistical analysis.** For statistical analyses, unpaired, two-tailed Student's  $t$ -tests (Excel) or one-way ANOVA analyses followed by a Dunnett's multiple comparison post-test were used as indicated in the respective figure legends. No statistical methods were used to predetermine sample size.

29. van Heerebeek, L. *et al.* Myocardial structure and function differ in systolic and diastolic heart failure. *Circulation* **113**, 1966–1973 (2006).
30. Kassiri, Z. *et al.* Combination of tumor necrosis factor- $\alpha$  ablation and matrix metalloproteinase inhibition prevents heart failure after pressure overload in tissue inhibitor of metalloproteinase-3 knock-out mice. *Circ. Res.* **97**, 380–390 (2005).
31. Nemir, M. *et al.* The Notch pathway controls fibrotic and regenerative repair in the adult heart. *Eur. Heart J.* **21**, 2174–2185 (2014).
32. Haefliger, J. A. *et al.* Connexin43-dependent mechanism modulates renin secretion and hypertension. *J. Clin. Invest.* **116**, 405–413 (2006).
33. Kuroda, H., Kutner, R. H., Bazan, N. G. & Reiser, J. Simplified lentivirus vector production in protein-free media using polyethylenimine-mediated transfection. *J. Virol. Methods* **157**, 113–121 (2009).
34. Ventura, A. *et al.* Cre-lox-regulated conditional RNA interference from transgenes. *Proc. Natl Acad. Sci. USA* **101**, 10380–10385 (2004).
35. Huang, L. E., Gu, J., Schau, M. & Bunn, H. F. Regulation of hypoxia-inducible factor 1 $\alpha$  is mediated by an O $_2$ -dependent degradation domain via the ubiquitin-proteasome pathway. *Proc. Natl Acad. Sci. USA* **95**, 7987–7992 (1998).
36. Troilo, A. *et al.* HIF1 $\alpha$  deubiquitination by USP8 is essential for cilogenesis in normoxia. *EMBO Rep.* **15**, 77–85 (2014).
37. Casonato, A. *et al.* A new L1446P mutation is responsible for impaired von Willebrand factor synthesis, structure, and function. *J. Lab. Clin. Med.* **144**, 254–259 (2004).
38. Krishnan, J. *et al.* Dietary obesity-associated Hif1 activation in adipocytes restricts fatty acid oxidation and energy expenditure via suppression of the Sirt2-NAD $^+$  system. *Genes Dev.* **26**, 259–270 (2012).
39. König, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Struct. Mol. Biol.* **17**, 909–915 (2010).
40. Hirschy, A., Schatzmann, F., Ehlers, E. & Perriard, J. C. Establishment of cardiac cytoarchitecture in the developing mouse heart. *Dev. Biol.* **289**, 430–441 (2006).
41. Ishimoto, T. *et al.* Opposing effects of fructokinase C and A isoforms on fructose-induced metabolic syndrome in mice. *Proc. Natl Acad. Sci. USA* **109**, 4320–4325 (2012).
42. Buescher, J. M., Moco, S., Sauer, U. & Zamboni, N. Ultrahigh performance liquid chromatography-tandem mass spectrometry method for fast and robust quantification of anionic and aromatic metabolites. *Anal. Chem.* **82**, 4403–4412 (2010).
43. Fuhrer, T., Heer, D., Begemann, B. & Zamboni, N. High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. *Anal. Chem.* **83**, 7074–7080 (2011).
44. Fukuzawa, J. *et al.* Cardiotrophin-1 increases angiotensinogen mRNA in rat cardiac myocytes through STAT3: an autocrine loop for hypertrophy. *Hypertension* **35**, 1191–1196 (2000).
45. Son, N.-H. *et al.* Cardiomyocyte aldose reductase causes heart failure and impairs recovery from ischemia. *PLoS ONE* **7**, e46549 (2012).
46. Radin, N. S. Extraction of tissue lipids with a solvent of low toxicity. *Methods Enzymol.* **72**, 5–7 (1981).
47. Koopman, R., Schaart, G. & Hesselink, M. K. C. Optimisation of oil red O staining permits combination with immunofluorescence and automated quantification of lipids. *Histochem. Cell Biol.* **116**, 63–68 (2001).
48. Kim, K.-Y. *et al.* Parkin is a lipid-responsive regulator of fat uptake in mice and mutant human cells. *J. Clin. Invest.* **121**, 3701–3712 (2011).

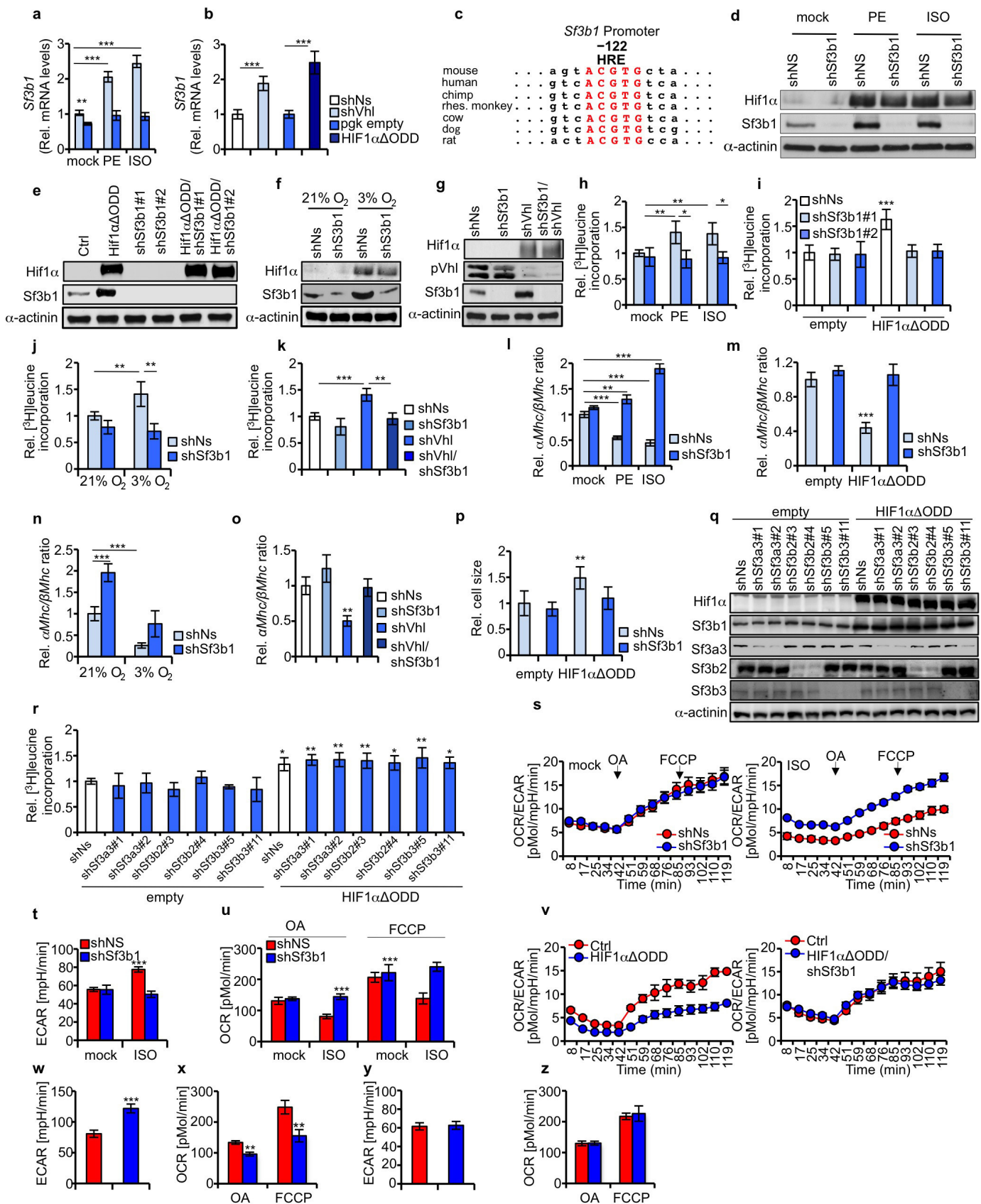




**Extended Data Figure 1 | HIF1 $\alpha$  induces SF3B1 gene expression that mediates KHK-C isoform expression.** a, Schematic experimental workflow.

b, Comparison of *C1QBP*, *HNRNP3*, *JMJD6* and *SF3B1* promoter sequences of human and mouse illustrate putative HREs up to 1,000 bp upstream of the transcription start site (TSS). The core consensus HRE motif is shown in colour. c, d, qPCR (c) and immunoblotting (d) for *Sf3b1* expression in NMCs transduced with scrambled shNs or viruses bearing two distinct shRNAs against *Sf3b1*. All values obtained by qPCR are presented in relation to *Sf3b1* expression in shNs (set as 1.0) ( $n = 3$  biological replicates per group). e, RT-PCR analysis of *Khk* mRNA as in Fig. 1 g in NMCs transduced as indicated. f, RT-PCR analyses of *Khk* mRNA from NMCs infected as indicated and stimulated with mock or isoproterenol (ISO). PCR products were digested with HincII as in Fig. 1 c. Splicing is quantified as percent of the *Khk-C* isoform (%*Khk-C*). g, RT-PCR based assessment of *Khk-A* and *Khk-C* isoform expression as done in Fig. 1 g. h, i, Ventricular samples of sham- or 1K1C-operated mice were analysed for *Khk* isoform expression (h) as done in Fig. 1 g or *Sf3b1*, *Khk-A* and *Khk-C* mRNA expression (i) by qPCR;  $n = 6$  for each group. j, k, Ventricular samples of sham- or TAC-treated mice (j) and NaCl- or

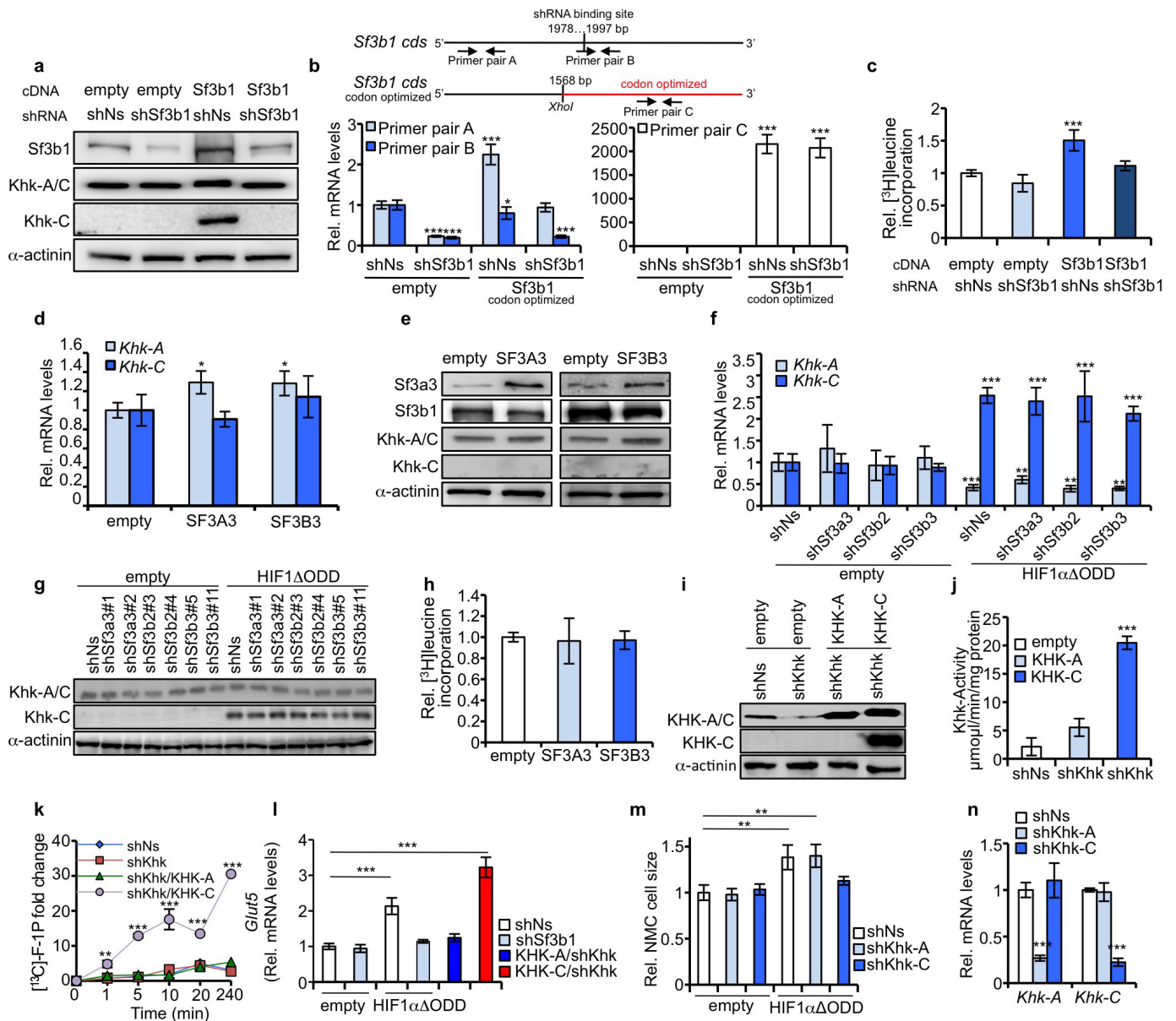
isoproterenol-treated mice (k) were analysed for expression of *Sf3b1*, *Khk-A* and *Khk-C* mRNA expression by qPCR ( $n = 6$  for each individual group). l, NMCs transduced as specified were validated for protein expression of indicated proteins by using immunoblots. In addition lysates of NMCs overexpressing *KHK-A* or *KHK-C* were processed for immunoblotting *Khk* antibodies. m, n, Ventricular samples of sham- or TAC-treated (m) and sham- or isoproterenol-treated mice (n) were analysed for indicated proteins by immunoblotting. o, Probes from left ventricles of 1K1C-, TAC- or isoproterenol-treated mice and their corresponding controls were analysed for mRNA expression of *Hif1α* targets. All values were presented in relation to sham-operated controls set as 1.0 ( $n = 6$  for each individual group). p, q, Left-ventricular biopsies of HCM patients and healthy controls were analysed for mRNA expression of *SF3B1*, *KHK-A* and *KHK-C* (p) and *HIF1α* targets (q). All values are presented in relation to healthy controls set as 1.0 ( $n = 6$  for controls and  $n = 16$  for patient samples). Error bars are s.d. (c) or s.e.m. (i-k, o-q). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; two-tailed unpaired *t*-test (c, i-k, o-q).



**Extended Data Figure 2 | *Sf3b1* is a HIF1 $\alpha$ -target gene, whose product regulates growth and metabolic shift *in vitro*.** **a**, NMCs infected as denoted, were stimulated with either mock, phenylephrine and isoproterenol, respectively. *Sf3b1* mRNA expression was examined by qPCR. ( $n = 3$  biological replicates). **b**, NMCs were transduced as indicated and processed for qPCR analysis of *Sf3b1* mRNA. ( $n = 3$  biological replicates). **c**, Comparative analysis of a conserved HRE (shown in red) in the promoter of *Sf3b1* of diverse mammalian species located 122 bp downstream of the transcription start site (TSS). **d–g**, Immunoblot detection of indicated proteins of NMCs transduced with shNs, and shSf3b lentiviruses treated with phenylephrine or isoproterenol (**d**) co-infected with empty vector or HIF1 $\alpha$ ΔODD (**e**) cultured under normoxic (21% O<sub>2</sub>) or hypoxic (3% O<sub>2</sub>) conditions (**f**) and co-transduced with shVhl lentiviruses (**g**). **h–o**, Evaluation of [<sup>3</sup>H]leucine incorporation and the myosin heavy chain  $\alpha/\beta$  mRNA expression ratio in NMCs treated with either phenylephrine or isoproterenol (**h**, **i**), HIF1 $\alpha$ ΔODD overexpression (**i**, **m**), hypoxia (**j**, **n**) or pVhl deletion (**k**, **o**), with or without shSf3b1 mediated Sf3b1 depletion. Obtained values are shown relative to incorporated

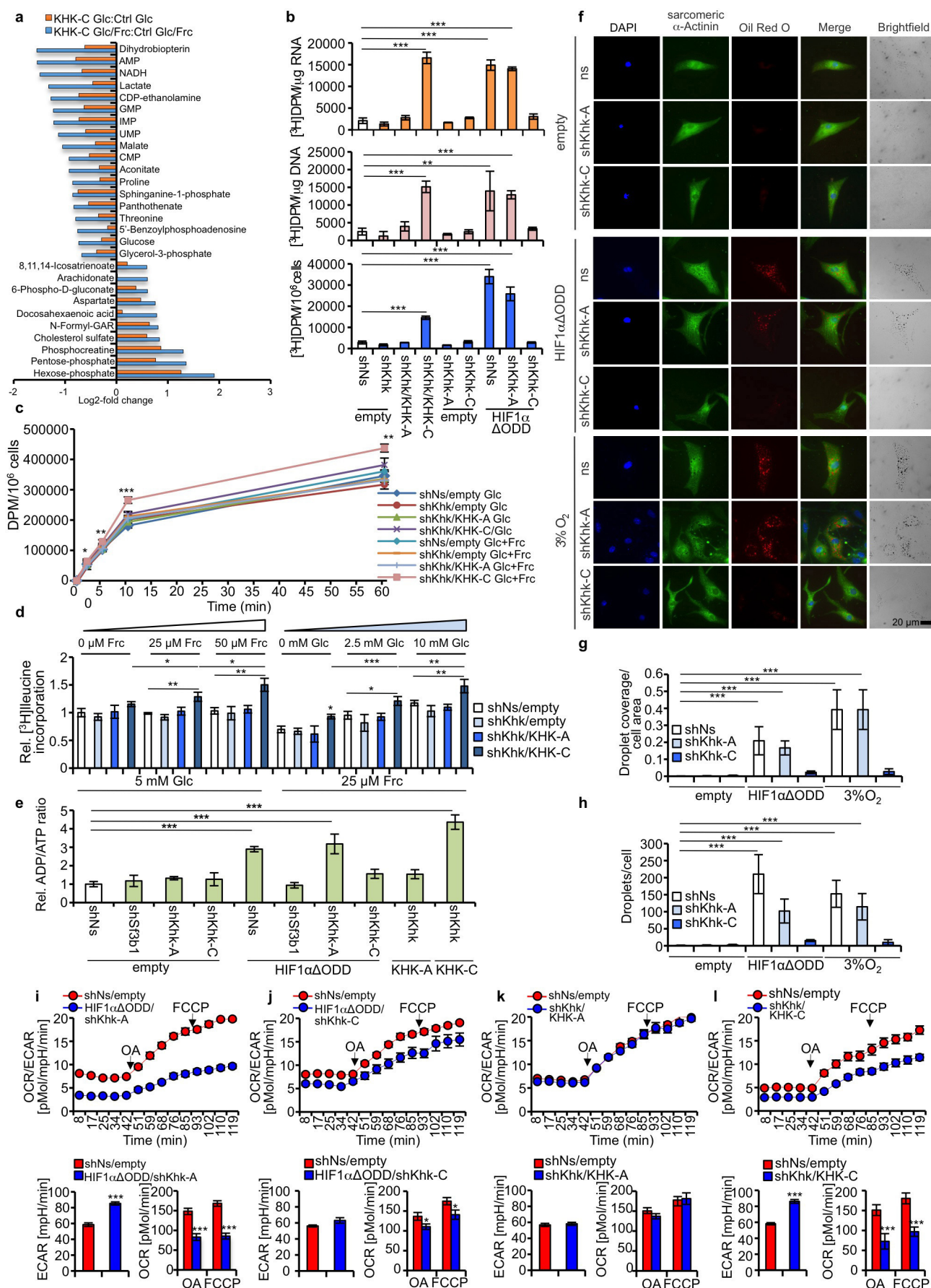
radioactivity in control (shNS) NMCs (set as 1.0) ( $n = 4$  per individual group). **p**, NMCs treated as in **i** were microscopically analysed for cell size using CellProfiler. Data represents 1 of 3 independent experiments with approximately 200 analysed cells per experiment. **q**, Lysates of NMCs treated as indicated, were assessed for protein expression of specified proteins by immunoblots. **r**, Evaluation of [<sup>3</sup>H]leucine incorporation of NMCs infected as indicated. Obtained values are shown relative to incorporated radioactivity in control (shNs/empty) NMCs (set as 1.0) ( $n = 3$  per individual group). **s–z**, Assessment of extracellular acidification rate (ECAR), and oleic acid- and carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone (FCCP)-induced oxygen consumption rate (OCR) in NMCs transduced and treated as indicated. Highlighted ECAR measurements were done at baseline. All data are compared to shNs-transduced NMCs. ( $n = 8$  biological replicates in the group shNs + isoproterenol;  $n = 10$  biological replicates in all other groups). Error bars are s.d. (**a**, **b**, **h–o**, **r**) or s.e.m. (**p**, **s–z**). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Two-tailed unpaired *t*-test (**b**, **t**, **u**, **w–z**). One-way ANOVA followed by Dunnett's multiple comparison post-test (**a**, **h–p**, **r**).





**Extended Data Figure 3 | Influence of U2snRNP-complex members on KHK isoform expression and validation of KHK-C-mediated fructose uptake and metabolism.** **a**, NMCs were transfected with either shNs or shSf3b1 and transfected with codon optimized Sf3b1 or empty construct and expression of denoted proteins was assessed. **b**, NMCs were transfected with either shNs or shSf3b1 and transfected with codon-optimized Sf3b1 or empty construct, respectively. Exogenous *Sf3b1*-expression was evaluated by qRT-PCR. Three different primer pairs were used. Primer pair A detects both endogenous and exogenous codon-optimized Sf3b1 (top). Primer pair B recognizes Sf3b1 close to the shRNA binding site and is specific for endogenous Sf3b1 (left). Primer pair C is located in the codon-optimized region (right). Values were normalized to shNs/empty-vector-infected NMCs ( $n = 3$  biological replicates per group). **c**, [ $^3$ H]leucine incorporation normalized to cell number in NMCs treated as in **b** ( $n = 4$  biological replicates per group). **d**, NMCs transfected with empty vector or human SF3A3 or SF3B3 were assessed for mRNA expression of *Khk-A* and *Khk-C* by qRT-PCR. Values were normalized to empty-vector-transduced NMCs ( $n = 3$  biological replicates per group). **e**, NMCs transfected as in **d** were immunoblotted for indicated proteins. **f**, NMCs transfected as indicated were probed for mRNA expression of *Khk-A* or *Khk-C* by qRT-PCR. Values were normalized to shNs/empty-vector-transduced NMCs ( $n = 3$  biological replicates). **g**, NMCs treated with either scrambled shRNA or two different shRNAs targeting either Sf3a3, Sf3b2

or Sf3b3 were co-transduced with empty vector or HIF1 $\Delta$ ODD. Protein expression of *Khk-A/C* or *Khk-C* specifically was evaluated by immunoblotting. **h**, Evaluation of [ $^3$ H]leucine incorporation of NMCs infected as in **d** normalized to empty-vector-transduced NMCs ( $n = 4$  biological replicates per group). **i**, Lysates of NMCs treated as indicated were immunoblotted for specified proteins. **j**, KHK-activity in lysates of NMCs transfected with a combination of lentiviruses expressing either shKhk targeting mouse *Khk* and human *KHK-A* or shKhk and human *KHK-C* or a scrambled shRNA control together with an empty vector. **k**, Evaluation of the kinetics of [ $^3$ C]fructose-1-phosphate accumulation in NMCs treated as in **i** at the respective time points ( $n = 3$  biological replicates). **l**, *Glut5* mRNA expression in NMCs transfected as indicated. ( $n = 3$  biological replicates per group). **m**, NMCs transfected as indicated were microscopically analysed for cell size. Data represents 1 of 3 experiments with approximately 50 analysed cells per experiment. **n**, *Khk-A*- and *Khk-C*-specific shRNAs were evaluated for their potential to inhibit *Khk-A* and *Khk-C* isoform expression, respectively. NMCs were infected with shNs, shKHK-A or shKHK-C lentiviruses. Data are shown in relation to shNs NMCs (set to 1.0) ( $n = 3$  biological replicates). Error bars are s.d. (**b–d**, **f**, **h**, **k**, **l**, **n**) or s.e.m. (**m**). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . One-way ANOVA followed by Dunnett's multiple comparison post-test (**b–d**, **f**, **h**, **k**, **l**, **m**, **n**).

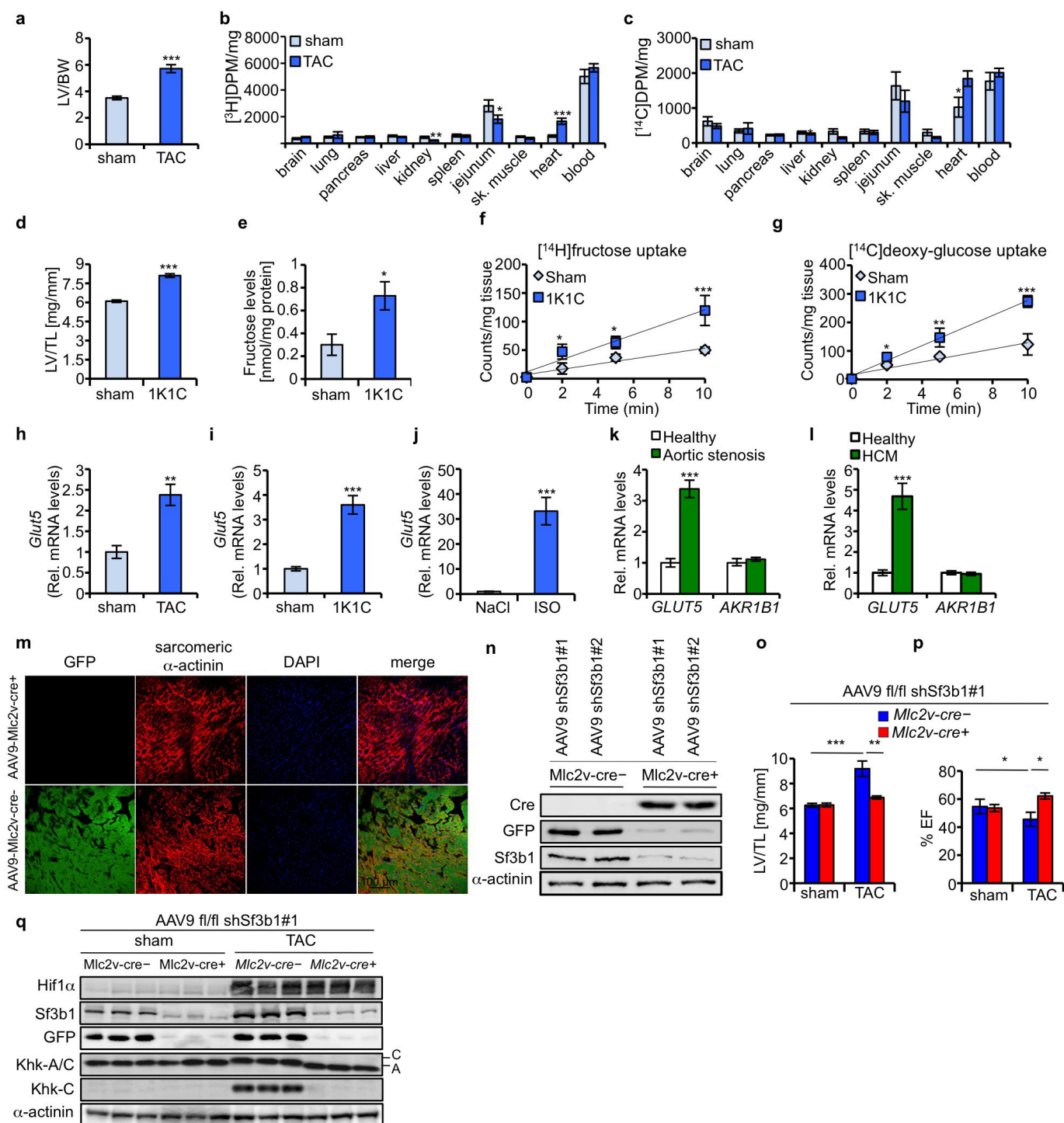


**Extended Data Figure 4 | Downstream metabolic effects of fructose**

**metabolism in NMCs.** **a**,  $\log_2$  fold change of metabolites in NMCs overexpressing KHK-C and cultured with glucose or glucose and fructose compared to the corresponding control transduced NMCs ( $n = 4$  biological replicates per group). **b**, Incorporation of [ $^3\text{H}$ ]fructose into RNA (top), DNA (middle) and protein (bottom) of NMCs treated as indicated ( $n = 4$  biological replicates per group). **c**, Uptake of [ $^{14}\text{C}$ ]deoxyglucose in NMCs infected as in **b** at depicted time points ( $n = 4$  biological replicates per group). **d**, [ $^3\text{H}$ ]leucine incorporation in NMCs transduced with shNs or shKHK and co-transduced with either empty overexpression vector or KHK-A or KHK-C respectively. NMCs were cultured in media with increasing fructose concentrations, under physiologic glucose amounts or with increasing glucose concentrations under physiologic fructose amounts. Data are presented relative to shNs/empty-vector-transduced NMCs at 5 mM glucose/0  $\mu\text{M}$  fructose ( $n = 4$  biological replicates). **e**, ADP/ATP ratios in NMCs transduced as indicated. ( $n = 4$  biological replicates, data show 1 of 3 representative experiments).

**f**, Immunofluorescence images of NMCs transduced with empty vector (upper panel) or HIF1 $\alpha$  $\Delta\text{ODD}$  (middle panel) lentiviruses or exposed to hypoxia (lower panel) were additionally transduced as indicated. Prior to staining for sarcomeric  $\alpha$ -actinin, Oil Red O and DAPI, NMCs were incubated with oleic acid. **g**, Quantification of lipid droplets/cell in NMCs of immunofluorescent images shown in **f**. **h**, Ratio of lipid droplets/cell area of NMCs shown in **g**. **i–l**, NMCs were infected as indicated and processed for oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) measurements under basal conditions or after injection of OA or FCCP. Depicted are rates expressed as OCR to ECAR ratios (upper panels) or individual ECAR (for measurements at baseline) or OCR values ( $n = 8$  biological replicates for HIF1 $\alpha$  $\Delta\text{ODD}$ /shKHK-A;  $n = 10$  biological replicates for all other groups). Error bars are s.d. (**b–e**) or s.e.m. (**g–l**). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . One-way ANOVA followed by Dunnett's multiple comparison post-test (**b–e**, **g**, **h**). Two-tailed unpaired  $t$ -test (**i–l**).





**Extended Data Figure 5 | Fructose and glucose uptake *ex vivo* and**

**validation of AAV9-fl/shSf3b1 viruses.** **a**, Ratio of left-ventricular weight to body weight of sham- or TAC-treated C57/Bl6J mice; ( $n = 7$  per group).

**b, c**, [ $^3\text{H}$ ]fructose (**b**) and [ $^{14}\text{C}$ ]deoxy-glucose (**c**) were co-applied by oral gavage and biodistribution was measured at 45 min post injection ( $n = 7$  per group).

**d**, Ratio of left-ventricular weight to tibia length (LV/TL) of sham- or 1K1C-treated C57/Bl6J wild-type mice ( $n = 6$  for sham and  $n = 7$  for 1K1C-treated mice).

**e**, Fructose levels in probes from left ventricles of indicated mice as determined by a quantitative colorimetric assay, normalized to protein amount ( $n = 4$  for sham and  $n = 5$  for hearts of 1K1C-treated mice).

**f, g**, Freshly isolated left-ventricular pieces of sham- or 1K1C-treated mice incubated in medium containing either [ $^{14}\text{C}$ ]fructose (**f**) or [ $^{14}\text{C}$ ]deoxyglucose (**g**) as indicated were processed for uptake measurements. Counts were normalized to tissue weight ( $n = 6$  per time point and treatment).

**h–j**, qPCR analyses of *Glut5* mRNA expression in left ventricles of 1K1C- (**h**), TAC- (**i**) or isoproterenol-treated (**j**) mice. All values are presented in relation to *Glut5* expression in shams (set as 1.0).  $n = 5$  for sham-treated mice as controls for TAC-treated mice ( $n = 6$ ), all other groups  $n = 7$ .

**k, l**, Relative mRNA expression of *GLUT5*

and aldose reductase (*AKR1B1*) in left-ventricular biopsies of patients diagnosed with aortic stenosis (**k**) or HCM (**l**) in relation to healthy controls (set as 1.0) was analysed by qPCR ( $n = 6$  for controls;  $n = 16$  for HCM and  $n = 17$  for patients with aortic stenosis).

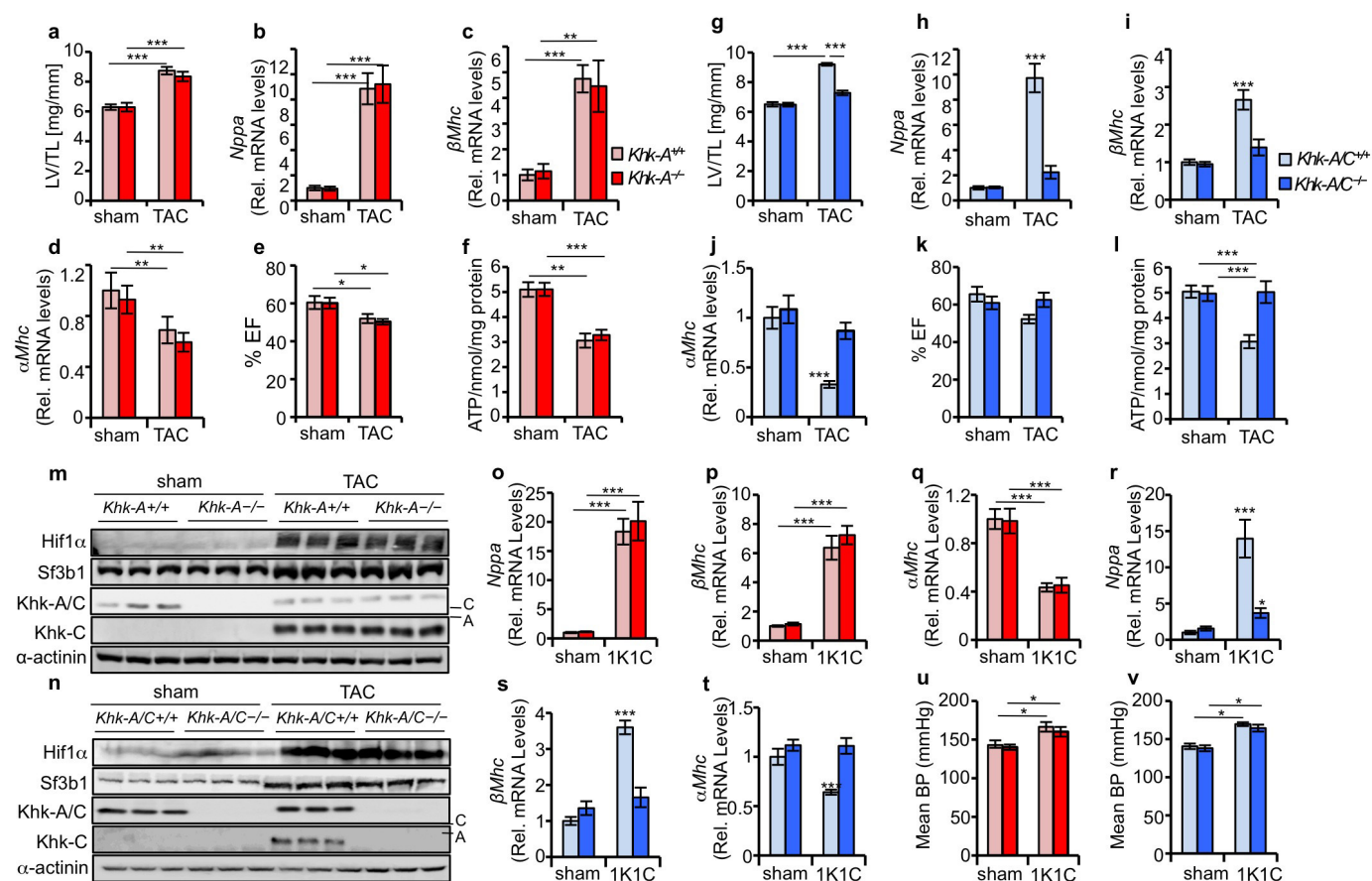
**m**, Integration of AAV9-fl/shSf3b1 viruses in the myocardium of *Mlc2-cre*<sup>−</sup> or *Mlc2v-cre*<sup>+</sup> mice was examined 12 weeks after intravenous injection by confocal microscopy (using 20× magnification) of left-ventricular cryoslices stained against GFP and sarcomeric  $\alpha$ -actinin and counterstained for DAPI.

**n**, Ventricular samples of *Mlc2v-cre*<sup>−</sup> or *Mlc2-cre*<sup>+</sup> mice transduced with AAV9-fl/shSf3b1#1 or AAV9-fl/shSf3b1#2 viruses were assessed for expression of indicated proteins by immunoblotting.

**o, p**, Ratio of left-ventricular weight to tibia length (**o**) and ejection fraction (%EF) (**p**) of sham- or TAC-treated *Mlc2v-cre*<sup>−</sup> and *Mlc2v-cre*<sup>+</sup> mice injected with AAV9-fl/shSf3b1#1 (number of mice per group is given in Extended Data Table 1).

**q**, Lysates of hearts from sham- or TAC-operated *Mlc2v-cre*<sup>−</sup> and *Mlc2v-cre*<sup>+</sup> mice transduced with AAV9-fl/shSf3b1#1 viruses were processed for immunoblotting with antibodies as denoted. Error bars are s.e.m. (**a–l, o, p**). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , two-tailed unpaired *t*-test (**a–l, o, p**).

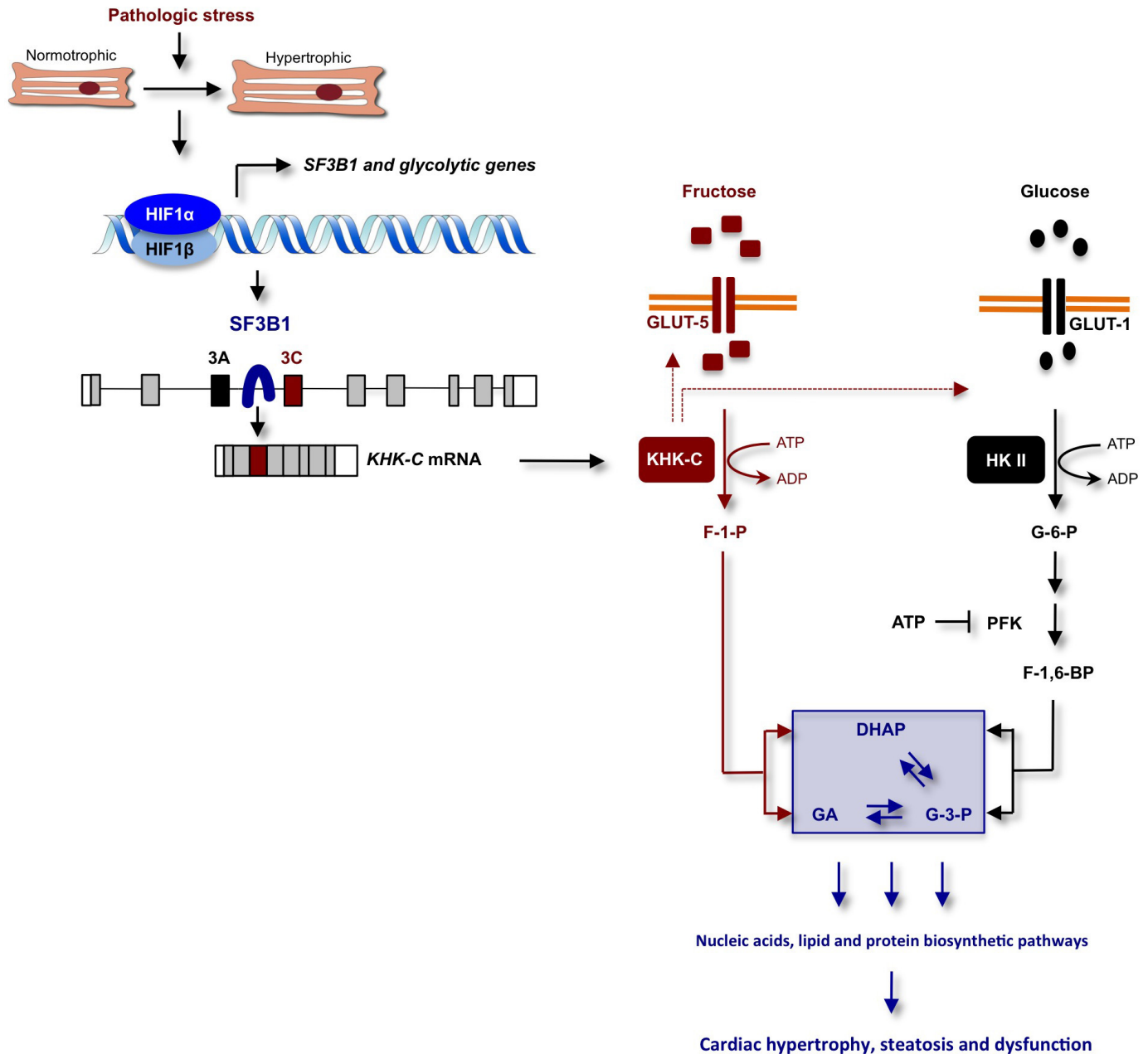
and aldose reductase (*AKR1B1*) in left-ventricular biopsies of patients diagnosed with aortic stenosis (**k**) or HCM (**l**) in relation to healthy controls (set as 1.0) was analysed by qPCR ( $n = 6$  for controls;  $n = 16$  for HCM and  $n = 17$  for patients with aortic stenosis). **m**, Integration of AAV9-fl/shSf3b1 viruses in the myocardium of *Mlc2-cre*<sup>−</sup> or *Mlc2v-cre*<sup>+</sup> mice was examined 12 weeks after intravenous injection by confocal microscopy (using 20× magnification) of left-ventricular cryoslices stained against GFP and sarcomeric  $\alpha$ -actinin and counterstained for DAPI. **n**, Ventricular samples of *Mlc2v-cre*<sup>−</sup> or *Mlc2-cre*<sup>+</sup> mice transduced with AAV9-fl/shSf3b1#1 or AAV9-fl/shSf3b1#2 viruses were assessed for expression of indicated proteins by immunoblotting. **o, p**, Ratio of left-ventricular weight to tibia length (**o**) and ejection fraction (%EF) (**p**) of sham- or TAC-treated *Mlc2v-cre*<sup>−</sup> and *Mlc2v-cre*<sup>+</sup> mice injected with AAV9-fl/shSf3b1#1 (number of mice per group is given in Extended Data Table 1). **q**, Lysates of hearts from sham- or TAC-operated *Mlc2v-cre*<sup>−</sup> and *Mlc2v-cre*<sup>+</sup> mice transduced with AAV9-fl/shSf3b1#1 viruses were processed for immunoblotting with antibodies as denoted. Error bars are s.e.m. (**a–l, o, p**). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , two-tailed unpaired *t*-test (**a–l, o, p**).



**Extended Data Figure 6 | Pathologic growth is dependent on HIF1 $\alpha$ -driven and SF3B1-mediated KHK-C expression.** **a–f**, Ratio of left-ventricular weight to tibia length (LV/TL) (**a**), mRNA expression of the hypertrophic markers *Nppa* (**b**), myosin heavy chain  $\beta$  ( $\beta$ Mhc) (**c**) and myosin heavy chain  $\alpha$  ( $\alpha$ Mhc) (**d**), ejection fraction (%EF) (**e**) and left-ventricular ATP content (**f**) of sham- or TAC-treated *Khk-A*<sup>+/+</sup> and *Khk-A*<sup>-/-</sup> mice. **g–l**, Ratio of left-ventricular weight to tibia length (LV/TL) (**g**), mRNA expression of the hypertrophic markers *Nppa* (**h**), myosin heavy chain  $\beta$  ( $\beta$ Mhc) (**i**) and myosin heavy chain  $\alpha$  ( $\alpha$ Mhc) (**j**), ejection fraction (%EF) (**k**) and left-ventricular ATP content (**l**) of *Khk-A/C*<sup>+/+</sup> and *Khk-A/C*<sup>-/-</sup> mice. **m, n**, Lysates from ventricles of sham- or

TAC-operated *Khk-A*<sup>+/+</sup> and *Khk-A*<sup>-/-</sup> mice (**m**) and *Khk-A/C*<sup>+/+</sup> and *Khk-A/C*<sup>-/-</sup> mice (**n**) were processed for immunoblotting with antibodies as indicated. **o–t**, Assessment of hypertrophic marker gene expression (*Nppa*,  $\beta$ Mhc and  $\alpha$ Mhc) in ventricles of TAC- vs sham-treated *Khk-A*<sup>-/-</sup> mice and their corresponding *Khk-A*<sup>+/+</sup> controls (**o–q**) as well as *Khk-A/C*<sup>-/-</sup> mice and their corresponding *Khk-A/C*<sup>+/+</sup> controls (**r–t**). **u, v**, Mean blood pressure (BP) in mmHg of *Khk-A*<sup>+/+</sup> and *Khk-A*<sup>-/-</sup> (**u**) as well as *Khk-A/C*<sup>+/+</sup> and *Khk-A/C*<sup>-/-</sup> mice (**v**) 8 weeks after sham or 1K1C treatment. For **a–l, o–v**, the number of mice per group is given in Extended Data Table 2. Error bars are s.e.m. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; two-tailed unpaired *t*-test.





**Extended Data Figure 7 | Model explaining how HIF1 $\alpha$  activation of SF3B1-dependent splicing of KHK regulates fructose and glucose metabolism to promote cardiac hypertrophy in response to pathologic stress.** In this model, pathologic stress leads to increased expression of HIF1 $\alpha$  and HIF1 $\alpha$ -dependent activation of genes encoding glycolytic enzymes and the splicing factor *SF3B1*. *SF3B1*, in turn, assembles at the branch-point sequence upstream of exon 3C of *KHK* pre-mRNA leading to an inclusion of this exon and KHK-C protein production. This shift in isoform expression from KHK-A to KHK-C in response to HIF1 $\alpha$ -*SF3B1* pathway activation drives KHK-C-dependent fructose uptake via stimulation of GLUT5 expression, the conversion of fructose to fructose-1-phosphate (F1P) and contributes simultaneously to the activation of glucose uptake and metabolism through a yet-to-be-determined mechanism. The model holds that the

unrestrained conversion of fructose to F1P by KHK-C limits ATP levels, thereby alleviating potential allosteric inhibition of phosphofructokinase (PFK) by ATP to maintain a high glycolytic flux. F1P is further metabolized to dihydroxyacetone phosphate (DHAP) and glyceraldehyde (GA). While DHAP serves as a precursor for glycerol synthesis, GA can be further converted to glyceraldehyde-3-phosphate (G3P), a key glycolytic intermediate. G3P can be channelled into the non-oxidative pentose phosphate pathway (PPP) supporting nucleic and amino acid biosynthesis. This metabolic constellation, created by the activation of the HIF1 $\alpha$ -*SF3B1*-KHK-C axis, increases macromolecular biosynthetic capacity essential for hypertrophic growth, steatosis and cardiac dysfunction. HKII and F-1,6-BP denotes hexokinase II and fructose-1,6-bisphosphate, respectively.

**Extended Data Table 1 | Echocardiographic analysis of sham-, 1K1C- or TAC-treated *Mlc2v-cre*<sup>-/-cre</sup><sup>+</sup> mice injected with AAV9-fl/fl-shSf3b1 viruses.**

|             | <i>Mlc2v-cre</i> <sup>-</sup> |                | <i>Mlc2v-cre</i> <sup>+</sup> |             |
|-------------|-------------------------------|----------------|-------------------------------|-------------|
|             | AAV9 fl/fl shSf3b1#1          |                | AAV9 fl/fl shSf3b1#1          |             |
|             | sham                          | 1K1C           | sham                          | 1K1C        |
| n           | 4                             | 3              | 5                             | 3           |
| BW (g)      | 27.5 ± 0.6                    | 27.3 ± 0.7     | 27.3 ± 0.6                    | 27.8 ± 0.8  |
| IVS;d (mm)  | 0.71 ± 0.03                   | 1.00 ± 0.03**  | 0.70 ± 0.02                   | 0.80 ± 0.04 |
| IVS;s (mm)  | 0.98 ± 0.02                   | 1.30 ± 0.05**  | 1.04 ± 0.09                   | 1.08 ± 0.08 |
| LVID;d (mm) | 3.97 ± 0.08                   | 4.13 ± 0.10    | 3.85 ± 0.11                   | 4.09 ± 0.10 |
| LVID;s (mm) | 3.08 ± 0.10                   | 3.30 ± 0.11    | 2.86 ± 0.15                   | 2.98 ± 0.18 |
| LVPW;d (mm) | 0.69 ± 0.01                   | 0.85 ± 0.03**  | 0.67 ± 0.02                   | 0.82 ± 0.01 |
| LVPW;s (mm) | 0.91 ± 0.35                   | 1.08 ± 0.12    | 0.93 ± 0.09                   | 0.97 ± 0.04 |
| %FS         | 27.8 ± 1.9                    | 25.3 ± 2.1     | 29.8 ± 1.6                    | 27.1 ± 2.8  |
| E/A         | 2.21 ± 0.26                   | 0.80 ± 0.06*** | 2.17 ± 0.17                   | 1.73 ± 0.11 |

|                    | <i>Mlc2v-cre</i> <sup>-</sup> |                | <i>Mlc2v-cre</i> <sup>+</sup> |               |
|--------------------|-------------------------------|----------------|-------------------------------|---------------|
|                    | AAV9 fl/fl shSf3b1#1          |                | AAV9 fl/fl shSf3b1#1          |               |
|                    | sham                          | TAC            | sham                          | TAC           |
| n                  | 5                             | 4              | 5                             | 5             |
| BW (g)             | 27.4 ± 0.8                    | 28.0 ± 0.9     | 26.5 ± 0.5                    | 26.9 ± 0.3    |
| IVS;d (mm)         | 0.70 ± 0.03                   | 0.98 ± 0.03*** | 0.68 ± 0.02                   | 0.84 ± 0.04   |
| IVS;s (mm)         | 0.96 ± 0.06                   | 1.37 ± 0.04**  | 0.97 ± 0.05                   | 1.18 ± 0.05   |
| LVID;d (mm)        | 3.94 ± 0.06                   | 4.04 ± 0.09    | 3.82 ± 0.06                   | 3.68 ± 0.06   |
| LVID;s (mm)        | 2.89 ± 0.13                   | 2.80 ± 0.09    | 2.71 ± 0.13                   | 2.54 ± 0.11   |
| LVPW;d (mm)        | 0.68 ± 0.02                   | 0.98 ± 0.04*** | 0.70 ± 0.02                   | 0.82 ± 0.03   |
| LVPW;s (mm)        | 0.97 ± 0.06                   | 1.37 ± 0.10**  | 0.99 ± 0.07                   | 1.11 ± 0.03   |
| %FS                | 30.4 ± 2.3                    | 19.9 ± 3.9*    | 30.04 ± 3.8                   | 34.2 ± 1.3    |
| E/A                | 1.62 ± 0.09                   | 1.72 ± 0.23    | 1.68 ± 0.02                   | 1.74 ± 0.12   |
| Aortic Vel. (mm/s) | 923 ± 58                      | 3920 ± 305***  | 878 ± 72                      | 3670 ± 190*** |

|                    | <i>Mlc2v-cre</i> <sup>-</sup> |                | <i>Mlc2v-cre</i> <sup>+</sup> |               |
|--------------------|-------------------------------|----------------|-------------------------------|---------------|
|                    | AAV9 fl/fl shSf3b1#2          |                | AAV9 fl/fl shSf3b1#2          |               |
|                    | sham                          | TAC            | sham                          | TAC           |
| n                  | 5                             | 4              | 5                             | 5             |
| BW (g)             | 27.4 ± 0.8                    | 28.0 ± 0.9     | 26.5 ± 0.5                    | 26.9 ± 0.3    |
| IVS;d (mm)         | 0.70 ± 0.03                   | 0.98 ± 0.03*** | 0.68 ± 0.02                   | 0.84 ± 0.04   |
| IVS;s (mm)         | 0.96 ± 0.06                   | 1.37 ± 0.04**  | 0.97 ± 0.05                   | 1.18 ± 0.05   |
| LVID;d (mm)        | 3.94 ± 0.06                   | 4.04 ± 0.09    | 3.82 ± 0.06                   | 3.68 ± 0.06   |
| LVID;s (mm)        | 2.89 ± 0.13                   | 2.80 ± 0.09    | 2.71 ± 0.13                   | 2.54 ± 0.11   |
| LVPW;d (mm)        | 0.68 ± 0.02                   | 0.98 ± 0.04*** | 0.70 ± 0.02                   | 0.82 ± 0.03   |
| LVPW;s (mm)        | 0.97 ± 0.06                   | 1.37 ± 0.10**  | 0.99 ± 0.07                   | 1.11 ± 0.03   |
| %FS                | 30.4 ± 2.3                    | 19.9 ± 3.9*    | 30.04 ± 3.8                   | 34.2 ± 1.3    |
| E/A                | 1.62 ± 0.09                   | 1.72 ± 0.23    | 1.68 ± 0.02                   | 1.74 ± 0.12   |
| Aortic Vel. (mm/s) | 923 ± 58                      | 3920 ± 305***  | 878 ± 72                      | 3670 ± 190*** |

BW, body weight; IVS;d, interventricular septum diameter in diastole, IVS;s, interventricular septum diameter in systole; LVID;d, left-ventricular inner diameter in diastole; LVID;s, left-ventricular inner diameter in systole; LVPW;d, left-ventricular posterior wall diameter in diastole, LVPW;s, left-ventricular posterior wall diameter in systole; %FS, fractional shortening; E/A, ratio of early (E) to late (A) ventricular filling velocities; Aortic Vel. = aortic velocity. Values shown are mean ± s.e.m.; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; two-tailed unpaired *t*-test.

Extended Data Table 2 | Echocardiographic analysis of sham- or TAC-operated *Khk-A*<sup>+/+</sup>/*Khk-A*<sup>-/-</sup> and *Khk-A/C*<sup>+/+</sup>/*Khk-A/C*<sup>-/-</sup> mice.

|                    | <i>Khk-A</i> <sup>+/+</sup> |                | <i>Khk-A</i> <sup>-/-</sup> |                |
|--------------------|-----------------------------|----------------|-----------------------------|----------------|
|                    | sham                        | TAC            | sham                        | TAC            |
| n                  | 6                           | 6              | 5                           | 6              |
| BW (g)             | 25.6 ± 0.7                  | 27.3 ± 0.8     | 25.4 ± 0.7                  | 25.6 ± 0.7     |
| IVS;d (mm)         | 0.73 ± 0.01                 | 0.98 ± 0.02*** | 0.72 ± 0.01                 | 1.00 ± 0.03*** |
| IVS;s (mm)         | 1.03 ± 0.07                 | 1.34 ± 0.04**  | 1.12 ± 0.01                 | 1.36 ± 0.06**  |
| LVID;d (mm)        | 3.85 ± 0.08                 | 3.80 ± 0.06    | 3.75 ± 0.10                 | 3.73 ± 0.07    |
| LVID;s (mm)        | 2.71 ± 0.16                 | 2.65 ± 0.11    | 2.66 ± 0.20                 | 2.70 ± 0.10    |
| LVPW;d (mm)        | 0.72 ± 0.01                 | 0.99 ± 0.04*** | 0.72 ± 0.01                 | 0.99 ± 0.03*** |
| LVPW;s (mm)        | 1.04 ± 0.03                 | 1.35 ± 0.03*** | 1.05 ± 0.01                 | 1.32 ± 0.05*** |
| %FS                | 34.0 ± 2.1                  | 26.3 ± 1.5*    | 34.3 ± 1.7                  | 26.0 ± 1.2*    |
| E/A                | 1.38 ± 0.11                 | 2.20 ± 0.05**  | 1.97 ± 0.4                  | 1.90 ± 0.4     |
| Aortic Vel. (mm/s) | 938 ± 55                    | 4250 ± 275***  | 880 ± 68                    | 3700 ± 308***  |

|                    | <i>Khk-A/C</i> <sup>+/+</sup> |                | <i>Khk-A/C</i> <sup>-/-</sup> |               |
|--------------------|-------------------------------|----------------|-------------------------------|---------------|
|                    | sham                          | TAC            | sham                          | TAC           |
| n                  | 5                             | 6              | 6                             | 5             |
| BW (g)             | 25.8 ± 0.6                    | 24.9 ± 0.8     | 24.51 ± 0.6                   | 24.4 ± 0.6    |
| IVS;d (mm)         | 0.72 ± 0.01                   | 1.00 ± 0.05**  | 0.64 ± 0.04                   | 0.83 ± 0.06   |
| IVS;s (mm)         | 1.15 ± 0.01                   | 1.42 ± 0.07*   | 1.06 ± 0.07                   | 1.19 ± 0.13   |
| LVID;d (mm)        | 3.53 ± 0.13                   | 3.70 ± 0.23    | 3.65 ± 0.13                   | 3.67 ± 0.04   |
| LVID;s (mm)        | 2.28 ± 0.27                   | 2.61 ± 0.33    | 2.49 ± 0.17                   | 2.45 ± 0.10   |
| LVPW;d (mm)        | 0.69 ± 0.03                   | 1.00 ± 0.05*** | 0.67 ± 0.02                   | 0.81 ± 0.04*  |
| LVPW;s (mm)        | 1.12 ± 0.03                   | 1.25 ± 0.03*   | 1.00 ± 0.06                   | 1.11 ± 0.07   |
| %FS                | 35.5 ± 3.0                    | 27.5 ± 1.6*    | 32.1 ± 2.30                   | 33.4 ± 2.9    |
| E/A                | 1.51 ± 0.09                   | 1.56 ± 0.17    | 1.77 ± 0.10                   | 1.64 ± 0.21   |
| Aortic Vel. (mm/s) | 928 ± 120                     | 4020 ± 232***  | 937 ± 59                      | 3600 ± 292*** |

BW, body weight; IVS;d, interventricular septum diameter in diastole; IVS;s, interventricular septum diameter in systole; LVID;d, left-ventricular inner diameter in diastole; LVID;s, left-ventricular inner diameter in systole; LVPW;d, left-ventricular posterior wall diameter in diastole; LVPW;s, left-ventricular posterior wall diameter in systole; %FS, fractional shortening; E/A, ratio of early (E) to late (A) ventricular filling velocities; Aortic Vel. = aortic velocity. Values shown are mean ± s.e.m.; \**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001; two-tailed unpaired *t*-test.



**Extended Data Table 3 | Echocardiographic analysis of sham- or 1K1C-operated *Khk-A*<sup>+/+</sup>/*Khk-A*<sup>-/-</sup> and *Khk-A/C*<sup>+/+</sup>/*Khk-A/C*<sup>-/-</sup> mice.**

|             | <i>Khk-A</i> <sup>+/+</sup> |                | <i>Khk-A</i> <sup>-/-</sup> |               |
|-------------|-----------------------------|----------------|-----------------------------|---------------|
|             | sham                        | 1K1C           | sham                        | 1K1C          |
| n           | 5                           | 6              | 5                           | 6             |
| BW (g)      | 27.2 ± 0.7                  | 27.1 ± 0.6     | 28.2 ± 2.0                  | 29.7 ± 0.5    |
| IVS;d (mm)  | 0.75 ± 0.01                 | 0.94 ± 0.03*** | 0.72 ± 0.02                 | 1.01 ± 0.06** |
| IVS;s (mm)  | 1.08 ± 0.05                 | 1.31 ± 0.05**  | 1.00 ± 0.04                 | 1.31 ± 0.09*  |
| LVID;d (mm) | 4.02 ± 0.10                 | 3.93 ± 0.12    | 3.91 ± 0.18                 | 4.37 ± 0.32   |
| LVID;s (mm) | 3.02 ± 0.08                 | 3.03 ± 0.14    | 2.90 ± 0.20                 | 3.45 ± 0.40   |
| LVPW;d (mm) | 0.73 ± 0.02                 | 0.99 ± 0.04*** | 0.74 ± 0.02                 | 0.97 ± 0.04** |
| LVPW;s (mm) | 1.01 ± 0.05                 | 1.30 ± 0.05**  | 1.06 ± 0.05                 | 1.29 ± 0.06*  |
| %FS         | 27.1 ± 1.5                  | 22.6 ± 1.4*    | 28.8 ± 3.4                  | 19.2 ± 3.05*  |
| E/A         | 1.97 ± 0.17                 | 1.47 ± 0.10*   | 1.83 ± 0.55                 | 1.27 ± 0.14   |

|             | <i>Khk-A/C</i> <sup>+/+</sup> |                | <i>Khk-A/C</i> <sup>-/-</sup> |              |
|-------------|-------------------------------|----------------|-------------------------------|--------------|
|             | sham                          | 1K1C           | sham                          | 1K1C         |
| n           | 5                             | 6              | 6                             | 5            |
| BW (g)      | 25.8 ± 0.6                    | 24.9 ± 0.8     | 24.51 ± 0.6                   | 24.4 ± 0.6   |
| IVS;d (mm)  | 0.72 ± 0.01                   | 1.00 ± 0.05**  | 0.64 ± 0.04                   | 0.83 ± 0.06  |
| IVS;s (mm)  | 1.15 ± 0.01                   | 1.42 ± 0.07*   | 1.06 ± 0.07                   | 1.19 ± 0.13  |
| LVID;d (mm) | 3.53 ± 0.13                   | 3.70 ± 0.23    | 3.65 ± 0.13                   | 3.67 ± 0.04  |
| LVID;s (mm) | 2.28 ± 0.27                   | 2.61 ± 0.33    | 2.49 ± 0.17                   | 2.45 ± 0.10  |
| LVPW;d (mm) | 0.69 ± 0.03                   | 1.00 ± 0.05*** | 0.67 ± 0.02                   | 0.81 ± 0.04* |
| LVPW;s (mm) | 1.12 ± 0.03                   | 1.25 ± 0.03*   | 1.00 ± 0.06                   | 1.11 ± 0.07  |
| %FS         | 35.5 ± 3.0                    | 27.5 ± 1.6*    | 32.1 ± 2.30                   | 33.4 ± 2.9   |
| E/A         | 1.51 ± 0.09                   | 1.56 ± 0.17    | 1.77 ± 0.10                   | 1.64 ± 0.21  |

BW, body weight; IVS;d, interventricular septum diameter in diastole; IVS;s, interventricular septum diameter in systole; LVID;d, left-ventricular inner diameter in diastole; LVID;s, left-ventricular inner diameter in systole; LVPW;d, left-ventricular posterior wall diameter in diastole; LVPW;s, left-ventricular posterior wall diameter in systole; %FS, fractional shortening; E/A, ratio of early (E) to late (A) ventricular filling velocities; Aortic Vel. = aortic velocity. Values shown are mean ± s.e.m.; \**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001; two-tailed unpaired *t*-test.

# Atomic structure of the APC/C and its mechanism of protein ubiquitination

Leifu Chang<sup>1\*</sup>, Ziguang Zhang<sup>1\*</sup>, Jing Yang<sup>1</sup>, Stephen H. McLaughlin<sup>1</sup> & David Barford<sup>1</sup>

**The anaphase-promoting complex (APC/C) is a multimeric RING E3 ubiquitin ligase that controls chromosome segregation and mitotic exit. Its regulation by coactivator subunits, phosphorylation, the mitotic checkpoint complex and interphase early mitotic inhibitor 1 (Emi1) ensures the correct order and timing of distinct cell-cycle transitions. Here we use cryo-electron microscopy to determine atomic structures of APC/C-coactivator complexes with either Emi1 or a UbcH10-ubiquitin conjugate. These structures define the architecture of all APC/C subunits, the position of the catalytic module and explain how Emi1 mediates inhibition of the two E2s UbcH10 and Ube2S. Definition of Cdh1 interactions with the APC/C indicates how they are antagonized by Cdh1 phosphorylation. The structure of the APC/C with UbcH10-ubiquitin reveals insights into the initiating ubiquitination reaction. Our results provide a quantitative framework for the design of future experiments to investigate APC/C functions *in vivo*.**

The activities of the diverse proteins that orchestrate the sequential biochemical and morphological changes intrinsic to the cell division cycle are controlled through the integration of protein phosphorylation, proteolysis and changes in gene expression. Two cullin-RING E3 ubiquitin ligases, the APC/C and Skp1/Cul1/F-box (SCF), catalyse the ubiquitination of multiple cell-cycle proteins to regulate their proteasome-mediated proteolysis. By regulating the ordered degradation of substrates such as cyclins, securin, mitotic kinases, and microtubule motors and assembly factors, the APC/C controls sister chromatid segregation, cytokinesis and the initiation of chromosome duplication<sup>1–3</sup>.

The APC/C is a large assembly comprising 19 subunits<sup>4</sup>. Its activity depends on the association of one of two coactivator subunits (either Cdc20 or Cdh1) that specify substrate recognition and stimulate the catalytic activity of APC/C–E2 complexes<sup>4–7</sup>. Coactivators engage the APC/C through a conserved amino (N)-terminal C-box motif and a carboxy (C)-terminal Ile–Arg (IR) tail<sup>2,3</sup>. Cyclin-dependent kinase (CDK) phosphorylates the APC/C, promoting Cdc20 association<sup>8,9</sup>. In contrast, Cdh1 is negatively regulated by phosphorylation<sup>8,10,11</sup>.

APC/C activity is repressed by the spindle assembly checkpoint (SAC) to ensure accurate sister chromosome segregation<sup>3</sup>. The SAC effector, the mitotic checkpoint complex (MCC), inhibits mitotic APC/C<sup>Cdc20</sup> whereas Emi1 inhibits APC/C<sup>Cdh1</sup> during interphase<sup>12</sup>. Similar to the MCC, Emi1 blocks destruction box (D box) recognition by APC/C-coactivator complexes. However, Emi1 also antagonizes the two E2s, UbcH10 and Ube2S, that pair with the APC/C and which are responsible for catalysing ubiquitin chain initiation and elongation, respectively<sup>13–17</sup>.

A quantitative mechanistic understanding of the APC/C requires structural information to atomic resolution. We previously described the human APC/C<sup>Cdh1</sup>–substrate complex (APC/C<sup>Cdh1.Hsl1</sup>) from a cryo-electron microscopy (cryo-EM) reconstruction, allowing definition of the overall architecture of the multi-subunit assembly<sup>4</sup>. Still, the atomic structures of several large and most small APC/C subunits, and the regulatory N-terminal domain (NTD) of Cdh1, were unknown, which together with the flexibility of the catalytic module comprising the RING subunit Apc11 and the C-terminal domain of Apc2 (Apc2<sup>CTD</sup>) limited our understanding of the complex.

Here, we used recombinant APC/C<sup>Cdh1</sup> in complex with Emi1 (APC/C<sup>Cdh1.Emi1</sup>) to determine an atomic model of the APC/C to 3.6 Å resolution (Extended Data Figs 1 and 2 and Extended Data Table 1). This structure, together with an APC/C<sup>Cdh1</sup>–UbcH10 complex, reveals details of the initiating ubiquitination reaction and how Emi1 inhibits UbcH10 and Ube2S<sup>13,14</sup>.

## Atomic structure of APC/C<sup>Cdh1.Emi1</sup>

Fig. 1a, b shows two views of the atomic model of APC/C<sup>Cdh1.Emi1</sup> (Supplementary Video 1). A local resolution map indicates that most of the static regions of the complex are at ~3.2 Å resolution (Extended Data Fig. 1). Specific regions of the molecule, including the catalytic module and the Cdh1<sup>WD40</sup> domain are more mobile and visualized at lower resolution (Extended Data Fig. 1).

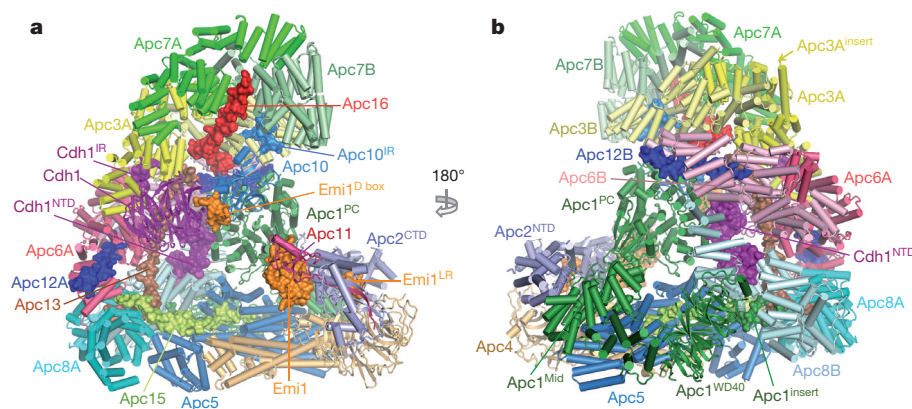
Similar to APC/C<sup>Cdh1.Hsl1</sup> (ref. 4), 60% of molecules are the coactivator-bound APC/C<sup>Cdh1.Emi1</sup> complex, with the remainder in the apo state (APC/C without coactivator). APC/C<sup>Cdh1.Emi1</sup> adopts the active conformation with the platform sub-domain, incorporating Apc2<sup>CTD</sup>–Apc11, in the upward position (Fig. 1a). However, in contrast with APC/C<sup>Cdh1.Hsl1</sup> where Apc2<sup>CTD</sup>–Apc11 is flexible<sup>4</sup>, Emi1 stabilizes Apc2<sup>CTD</sup>–Apc11 by bridging it to Apc1 (Fig. 1a).

With the APC/C<sup>Cdh1.Emi1</sup> EM map at 3.6 Å resolution, we built a complete atomic model of the complex. Existing crystal structures, complemented by newly determined structures of Apc4 and the N-terminal domain of Apc5 (unpublished data), together with homology models (Extended Data Tables 2 and 3), were used to guide building of most large APC/C subunits. For Apc1, fitting to the N-terminal Apc1<sup>WD40</sup> domain and the densities N- and C-terminal to its central PC domain (Apc1<sup>PC</sup>) were performed *ab initio*. These two regions coalesce to form Apc1<sup>Mid</sup> that connects Apc1<sup>WD40</sup> with Apc1<sup>PC</sup> and comprises an  $\alpha$ -solenoid capped by a  $\beta$ -sandwich (Extended Data Fig. 3a). Apc13 and Apc16 of the tetratricopeptide repeat (TPR) lobe interact with structurally homologous symmetry-related sites on the four TPR homo-dimers (Extended Data Fig. 3b, c). The refined APC/C<sup>Cdh1.Emi1</sup> model has excellent stereochemistry and is in good agreement with the EM density map (Extended Data Fig. 2 and Extended Data Table 1).

<sup>1</sup>MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK.

\*These authors contributed equally to this work.

**Figure 1 | EM reconstructions of the APC/C<sup>Cdh1,Emi1</sup> complex.** **a, b**, Two views of the atomic structure of APC/C<sup>Cdh1,Emi1</sup>. Large subunits are shown in cartoon, whereas the four small subunits, Emi1, Cdh1<sup>NTD</sup>, Cdh1<sup>IR</sup> and Apc10<sup>IR</sup>, are shown as surface representations. Emi1 interacts with the substrate-recognition and catalytic modules.

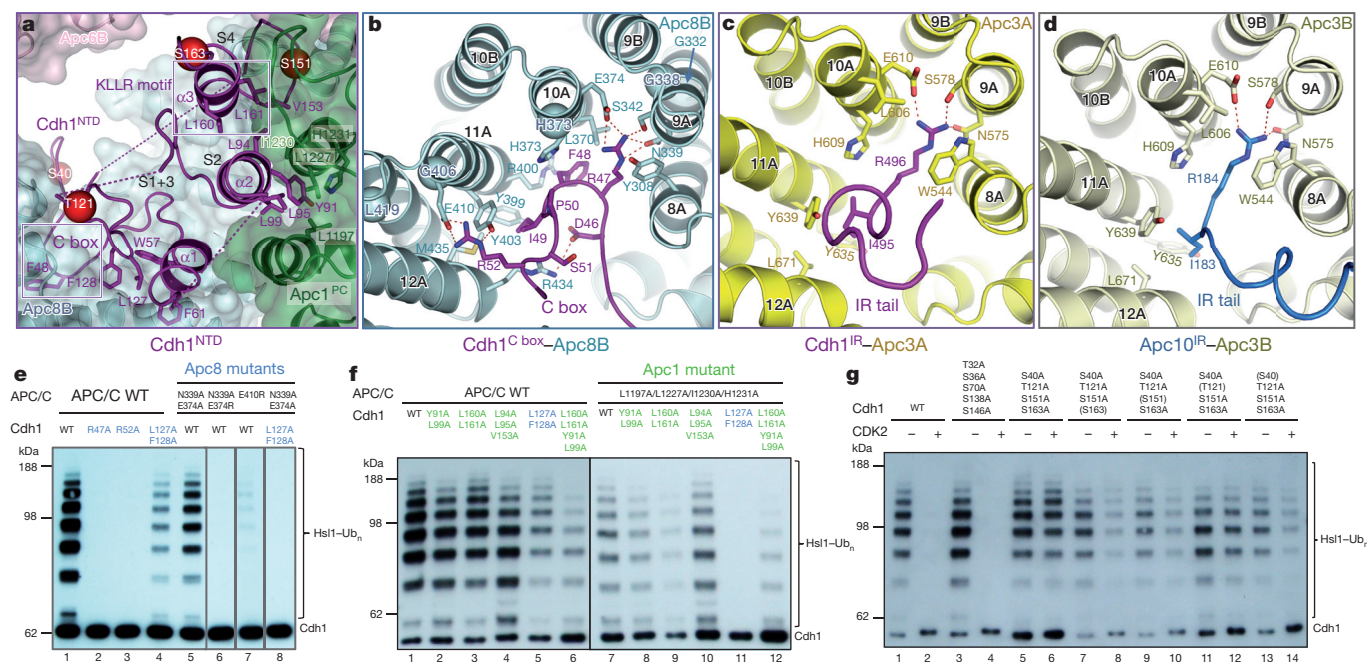


### Cdh1<sup>NTD</sup> binds to Apc1 and Apc8

In the APC/C<sup>Cdh1,Emi1</sup> EM map, the Cdh1<sup>NTD</sup>-assigned region shows clearly resolved side-chain features (Extended Data Fig. 1g). Four distinct segments of Cdh1<sup>NTD</sup> are organized into two mini-domains (Fig. 2a). Segments 1 and 3 interact with Apc8B, whereas segments 2 and 4 interact with Apc1<sup>PC</sup>. Segment 1 incorporates the seven-residue C box (residues 46–52) (Extended Data Fig. 3d), which, through extensive contacts with Apc8, is a key determinate of Cdh1<sup>NTD</sup>–APC/C interactions. The C box adopts a hairpin conformation that inserts into the TPR superhelix of Apc8 (Fig. 2a, b). Asp46, Arg47 and Arg52 of the C box form a network of electrostatic interactions, whereas Phe48 and Ile49 mediate non-polar contacts. Strikingly, Arg47 and Ile49 bind a site on Apc8 that is structurally homologous to the IR tail-binding site of Apc3, a paralogue of Apc8 (Fig. 2b–d). In both TPR proteins, three invariant residues (Asn339, Ser342 and Glu374 of Apc8) coordinate the Arg guanidinium group, with the aliphatic moiety of Arg sandwiched

between a conserved aromatic residue and an invariant Leu (Fig. 2b–d). The aliphatic side chain of Ile49 of the C box is flanked by two tyrosines (both conserved in Apc3) and a Met (Leu in Apc3). The main role of segment 3 is to stabilize the conformation of segment 1 by contributing Leu127 and Phe128 to a cluster of hydrophobic residues at the Apc8–Cdh1<sup>NTD</sup> interface (Fig. 2a). Mutation of both residues severely abrogates Cdh1 activity (Fig. 2e, lanes 1 and 4).

Segment 2 forms a leucine zipper-like  $\alpha$ -helical interface with two outer  $\alpha$ -helices of Apc1<sup>PC</sup>, an interaction mediated by three conserved leucines (Fig. 2a). The  $\alpha$ -helix of segment 4 aligns antiparallel to segment 2 (Fig. 2a). It includes the KLLR sequence (Extended Data Fig. 3d), structurally similar to the KLLR motif of Cdc20<sup>NTD</sup> that interacts with Mad2 as a  $\beta$ -strand<sup>18</sup>, and is implicated in mediating APC/C–Cdc20 interactions<sup>19</sup>. Leu160 and Leu161 of the KLLR sequence, together with Val153 of segment 4, contribute to the hydrophobic interface of segment 2 and Apc1<sup>PC</sup> (Fig. 2a, f).



**Figure 2 | The C box of Cdh1 and IR tails of Cdh1 and Apc10 interact with structurally related sites on Apc8 and Apc3.** **a**, Cdh1<sup>NTD</sup> binds to Apc8 and Apc1. Segments 1 and 3 of Cdh1<sup>NTD</sup> including the C box interact with Apc8, whereas segments 2 and 4 interact with Apc1<sup>PC</sup>. Segments 1–4 are labelled as S1–4. Phosphorylation sites Ser40, Thr121, Ser151 and Ser163 are shown as red spheres. **b**, Details of the C-box interactions. The Arg47 and Ile49 interaction sites are structurally related to the IR tail-binding sites for Cdh1 and Apc10 in Apc3 shown in **c** and **d**. *S. cerevisiae* CDC23 (Apc8) temperature-

sensitive mutations<sup>23</sup> map to the C-box binding site (C $\alpha$  of mutant residues shown as spheres in **b**). **e**, Mutation of C-box residues Arg47 and Arg52 eliminates APC/C<sup>Cdh1</sup> activity, as do mutations of Apc8 residues that interact with either Arg47 or Arg52. WT, wild type. Ub<sub>n</sub>, polyubiquitin. **f**, At the Cdh1<sup>NTD</sup>–Apc1<sup>PC</sup> interface, multiple residues cooperate to mediate APC/C–Cdh1 interactions. **g**, Ser40, Ser151 and Ser163 mediate the negative regulation of Cdh1 by CDK phosphorylation. Experiments in Fig. 2e, f were replicated three times and in Fig. 2g five times.





(Extended Data Fig. 2f). However, density corresponding to UbcH10, connected to  $\text{Apc2}^{\text{CTD}}\text{-Apc11}^{\text{RING}}$ , was weak.

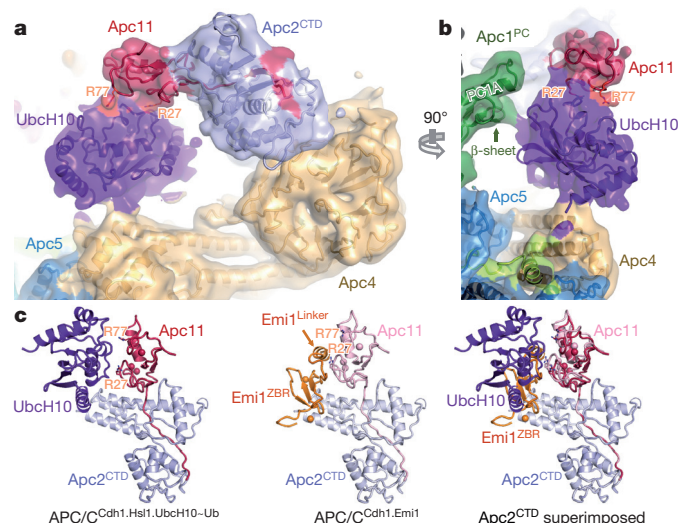
Extensive three-dimensional classification of APC/ $\text{C}^{\text{Cdh1.Hsl1.UbcH10-Ub}}$  revealed structural variability of the UbcH10 density consistent with low occupancy of the UBC domain (~36%), coupled with minor conformational variability of the  $\text{Apc2}^{\text{CTD}}\text{-Apc11-UbcH10}$  catalytic module (Extended Data Fig. 6a, cycle 2). Subjecting a major three-dimensional class for further classification generated a three-dimensional class with clearly resolved UbcH10 density (Extended Data Fig. 6a, cycle 3, class 5, and Extended Data Fig. 6b, c). This structural class was suitable for docking UbcH10 coordinates (Fig. 4a, b) and is representative of other APC/C-UbcH10 conformational states (Extended Data Fig. 6d, e). Relative to APC/C $^{\text{Cdh1.Emi1}}$ ,  $\text{Apc2}^{\text{CTD}}\text{-Apc11}^{\text{RING}}$  tilts slightly at the  $\text{Apc2}^{\text{NTD}}\text{-Apc2}^{\text{CTD}}$  interface (Extended Data Fig. 5f), and  $\text{Apc11}^{\text{RING}}$  tilts relative to  $\text{Apc2}^{\text{CTD}}$  (Fig. 4c). UbcH10 and  $\text{Apc11}^{\text{RING}}$  interact through a canonical E2-RING mechanism<sup>26</sup>, thus docking of UbcH10- $\text{Apc11}^{\text{RING}}$  was guided by crystal structures of RING-UbcH5-ubiquitin complexes<sup>29,30</sup> (Fig. 4a, b). Importantly, this showed that density for ubiquitin was not visible in the EM map, indicating that the ubiquitin moiety is flexible (Figs 4a, b and 5a). Consistent with this conclusion, a reconstruction of an APC/C $^{\text{Cdh1.Hsl1.Apc11-UbcH10}}$  fusion complex (Extended Data Figs 1b and 2d and Extended Data Table 1b), and a negative-stain structure of APC/C $^{\text{Cdh1.Hsl1}}$  with a huge excess of UbcH10, showed that the shape of the UbcH10-assigned density resembled the UbcH10-ubiquitin density in the APC/C $^{\text{Cdh1.Hsl1.UbcH10-Ub}}$  map (Extended Data Figs 6c and 7).

Since UbcH10 and  $\text{Apc11}^{\text{RING}}$  interact through a canonical E2-RING interface (Fig. 4a, b), the APC/C probably stimulates the intrinsic catalytic activity of UbcH10-ubiquitin<sup>31</sup> by promoting a closed conformation that primes the UbcH10-ubiquitin thioester bond<sup>15,29,30,32-36</sup>. In agreement with this, mutation of Ile36 that contacts  $\text{Apc11}^{\text{RING}}$  (refs 29, 30, 34) and Ile44 that contacts UbcH10 (refs 15, 29, 30, 32, 34) (Fig. 5a) ablates APC/C-catalysed ubiquitination (Extended Data Fig. 4e, f). Nevertheless, the absence of EM density for the ubiquitin moiety of the UbcH10<sup>LR</sup>-ubiquitin conjugate indicates that APC/C does not stabilize a static closed UbcH10-ubiquitin conformation<sup>29,30,34</sup>. Modelling shows that only  $\text{Apc11}^{\text{RING}}$  would interact with ubiquitin in the closed UbcH10-ubiquitin

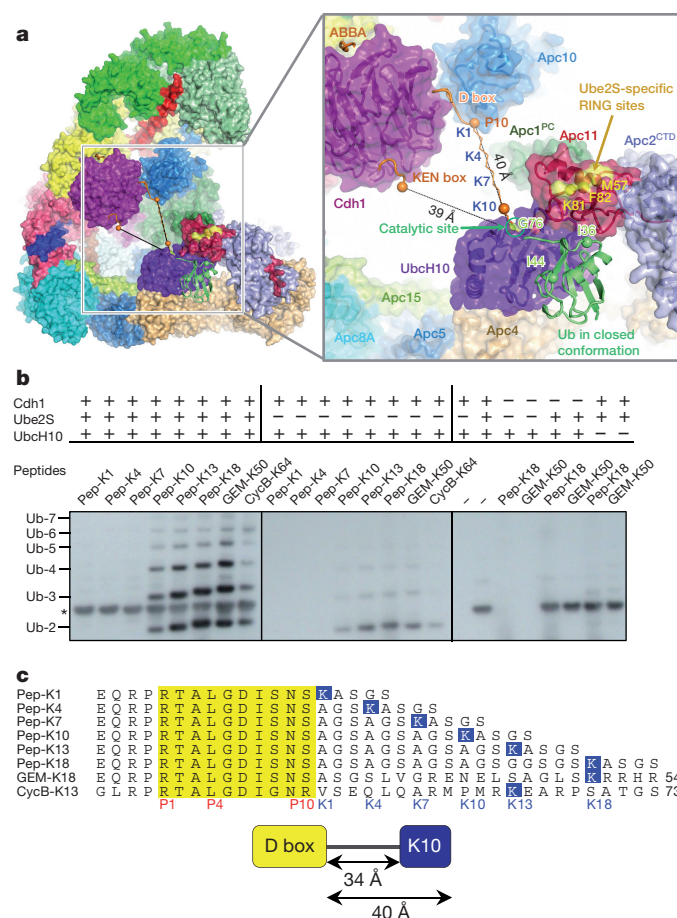
conformation (Fig. 5a), whereas in crystal structures of RING-UbcH5-ubiquitin conjugates where ubiquitin is static, either a second RING subunit<sup>29,30</sup> or a non-RING element<sup>34</sup> also contact ubiquitin to enhance catalytic efficiency. The APC/C is reminiscent of other single domain RING and U-box E3s that bias the E2-ubiquitin conformation from multiple extended states to the closed state<sup>33,37</sup>. An interesting possibility is that substrate initiation motifs that promote lysine ubiquitination<sup>38</sup> may induce a closed UbcH10-ubiquitin conformation.

## Defining target lysine ubiquitination

The APC/C $^{\text{Cdh1.Hsl1.UbcH10-Ub}}$  model shows that the UbcH10 catalytic Cys is some 40 Å from the C terminus of a D box at the D-box receptor<sup>39</sup> (Fig. 5a). An extended linker of ten residues would place a lysine C-terminal to the D box at the UbcH10-ubiquitin catalytic site (Fig. 5a). Consistent with this, peptides with a lysine at positions K10, K13 and K18 relative to the D box were ubiquitinated, but not peptides with lysines at K1, K4 and K7 (Fig. 5b, c). Peptides modelled on known



**Figure 4 | Structure of the APC/C $^{\text{Cdh1.Hsl1.UbcH10-Ub}}$  complex reveals the location of UbcH10. a, b, Views of the EM density at the catalytic centre with UbcH10-ubiquitin. No density for the ubiquitin moiety was recovered. Arg27 and Arg77 of Apc11, required for UbcH10 interactions<sup>26</sup>, are indicated. c, The UbcH10- and Emi1-binding sites of Apc11<sup>RING</sup> overlap. The Apc11<sup>RING</sup> orientation differs slightly in the two complexes.**



**Figure 5 | The relative position of the catalytic and substrate-recognition modules defines the target lysine. a, View of APC/C $^{\text{Cdh1.Hsl1.UbcH10-Ub}}$  with close-up view of the catalytic and substrate-recognition modules. A UbcH10-ubiquitin conjugate is modelled in the closed conformation<sup>29</sup>. The Ube2S-specific site of Apc11<sup>RING</sup> (for distal ubiquitin positioning)<sup>26</sup> is indicated. The KEN box and ABBA motif are modelled on ref. 39. The distance from the D and KEN boxes to the catalytic centre of UbcH10 (measured as C $\alpha$ -CO distance: P10 of D box and 'N' of KEN box to Gly76 of ubiquitin) is indicated. A ten-residue linker is modelled connected to P10 of the D box with Lys at K10. Spheres denote C $\alpha$ -atoms of D-box P10 and linker K10. b, The reactivity of defined peptides with progressively longer spacers separating the D box and acceptor lysine residue (sequences shown in c). Only peptides with a lysine ten or more residues C-terminal to the D box are substrates. \*Ube2S-specific band. The experiment was replicated three times. c, Peptide sequences used in ubiquitination assays. D box in yellow, target lysines in blue.**



APC/C substrates, geminin (Lys50, equivalent to K18 (ref. 38)) and cyclin B (Lys64, equivalent to K13), were also ubiquitinated.

## Conclusions

The atomic structure of the APC/C provides a quantitative framework for understanding APC/C mechanisms and the basis for future studies to investigate mechanisms of control by phosphorylation and the MCC. We speculate that in the APC/C<sup>Cdc20</sup> complex inhibited by MCC, the second C-box receptor of the Apc8 homo-dimer may bind the Cdc20 subunit of the MCC<sup>3,40</sup>. TAME is an Ile–Arg dipeptide APC/C inhibitor<sup>41</sup>. Its proposed mechanism is to inhibit coactivator association by competing for the IR tail-binding site on Apc3. The homology of the C-box receptor of Apc8 and the IR tail-receptor of Apc3 suggests that TAME may also disrupt C-box–APC/C interactions, a notion consistent with the more important role of the C box for APC/C activity and cell viability<sup>20</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 January; accepted 10 April 2015.

Published online 15 June 2015.

- Meyer, H. J. & Rape, M. Processive ubiquitin chain formation by the anaphase-promoting complex. *Semin. Cell Dev. Biol.* **22**, 544–550 (2011).
- Pines, J. Cubism and the cell cycle: the many faces of the APC/C. *Nature Rev. Mol. Cell Biol.* **12**, 427–438 (2011).
- Primorac, I. & Musacchio, A. Panta rhei: the APC/C at steady state. *J. Cell Biol.* **201**, 177–189 (2013).
- Chang, L., Zhang, Z., Yang, J., McLaughlin, S. H. & Barford, D. Molecular architecture and mechanism of the anaphase-promoting complex. *Nature* **513**, 388–393 (2014).
- Kimata, Y., Baxter, J. E., Fry, A. M. & Yamano, H. A role for the Fizzy/Cdc20 family of proteins in activation of the APC/C distinct from substrate recruitment. *Mol. Cell* **32**, 576–583 (2008).
- Van Voorhis, V. A. & Morgan, D. O. Activation of the APC/C ubiquitin ligase by enhanced E2 efficiency. *Curr. Biol.* **24**, 1556–1562 (2014).
- Kelly, A., Wickliffe, K. E., Song, L., Fedrigo, I. & Rape, M. Ubiquitin chain elongation requires E3-dependent tracking of the emerging conjugate. *Mol. Cell* **56**, 232–245 (2014).
- Kramer, E. R., Scheuringer, N., Podtelejnikov, A. V., Mann, M. & Peters, J. M. Mitotic regulation of the APC activator proteins CDC20 and CDH1. *Mol. Biol. Cell* **11**, 1555–1569 (2000).
- Rudner, A. D. & Murray, A. W. Phosphorylation by Cdc28 activates the Cdc20-dependent activity of the anaphase-promoting complex. *J. Cell Biol.* **149**, 1377–1390 (2000).
- Jaspersen, S. L., Charles, J. F. & Morgan, D. O. Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Curr. Biol.* **9**, 227–236 (1999).
- Zachariae, W., Schwab, M., Nasmyth, K. & Seufert, W. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science* **282**, 1721–1724 (1998).
- Reimann, J. D. et al. Emi1 is a mitotic regulator that interacts with Cdc20 and inhibits the anaphase promoting complex. *Cell* **105**, 645–655 (2001).
- Frye, J. J. et al. Electron microscopy structure of human APC/C<sup>CDH1</sup>–EMI1 reveals multimodal mechanism of E3 ligase shutdown. *Nature Struct. Mol. Biol.* **20**, 827–835 (2013).
- Wang, W. & Kirschner, M. W. Emi1 preferentially inhibits ubiquitin chain elongation by the anaphase-promoting complex. *Nature Cell Biol.* **15**, 797–806 (2013).
- Wickliffe, K. E., Lorenz, S., Wemmer, D. E., Kuriyan, J. & Rape, M. The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2. *Cell* **144**, 769–781 (2011).
- Williamson, A. et al. Identification of a physiological E2 module for the human anaphase-promoting complex. *Proc. Natl Acad. Sci. USA* **106**, 18213–18218 (2009).
- Wu, T. et al. UBE2S drives elongation of K11-linked ubiquitin chains by the anaphase-promoting complex. *Proc. Natl Acad. Sci. USA* **107**, 1355–1360 (2010).
- Sironi, L. et al. Crystal structure of the tetrameric Mad1–Mad2 core complex: implications of a ‘safety belt’ binding mechanism for the spindle checkpoint. *EMBO J.* **21**, 2496–2506 (2002).
- Izawa, D. & Pines, J. Mad2 and the APC/C compete for the same site on Cdc20 to ensure proper chromosome segregation. *J. Cell Biol.* **199**, 27–37 (2012).
- Thornton, B. R. et al. An architectural map of the anaphase-promoting complex. *Genes Dev.* **20**, 449–460 (2006).
- Matyskiela, M. E. & Morgan, D. O. Analysis of activator-binding sites on the APC/C supports a cooperative substrate-binding mechanism. *Mol. Cell* **34**, 68–80 (2009).
- Izawa, D. & Pines, J. How APC/C–Cdc20 changes its substrate specificity in mitosis. *Nature Cell Biol.* **13**, 223–233 (2011).
- Sikorski, R. S., Michaud, W. A. & Hieter, P. p62cdc23 of *Saccharomyces cerevisiae*: a nuclear tetratricopeptide repeat protein with two mutable domains. *Mol. Cell Biol.* **13**, 1212–1221 (1993).
- Lukas, C. et al. Accumulation of cyclin B1 requires E2F and cyclin-A-dependent rearrangement of the anaphase-promoting complex. *Nature* **401**, 815–818 (1999).
- Miller, J. J. et al. Emi1 stably binds and inhibits the anaphase-promoting complex/cyclosome as a pseudosubstrate inhibitor. *Genes Dev.* **20**, 2410–2420 (2006).
- Brown, N. G. et al. Mechanism of polyubiquitination by human anaphase-promoting complex: RING repurposing for ubiquitin chain assembly. *Mol. Cell* **56**, 246–260 (2014).
- Duda, D. M. et al. Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134**, 995–1006 (2008).
- Meyer, H. J. & Rape, M. Enhanced protein degradation by branched ubiquitin chains. *Cell* **157**, 910–921 (2014).
- Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012).
- Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7–E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nature Struct. Mol. Biol.* **19**, 876–883 (2012).
- Ozkan, E., Yu, H. & Deisenhofer, J. Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. *Proc. Natl Acad. Sci. USA* **102**, 18890–18895 (2005).
- Saha, A., Lewis, S., Kleiger, G., Kuhlman, B. & Deshaies, R. J. Essential role for ubiquitin-ubiquitin-conjugating enzyme interaction in ubiquitin discharge from Cdc34 to substrate. *Mol. Cell* **42**, 75–83 (2011).
- Pruneda, J. N. et al. Structure of an E3:E2–Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Mol. Cell* **47**, 933–942 (2012).
- Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. Essentiality of a non-RING element in priming donor ubiquitin for catalysis by a monomeric E3. *Nature Struct. Mol. Biol.* **20**, 982–986 (2013).
- Scott, D. C. et al. Structure of a RING E3 trapped in action reveals ligation mechanism for the ubiquitin-like protein NEDD8. *Cell* **157**, 1671–1684 (2014).
- Reverter, D. & Lima, C. D. Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. *Nature* **435**, 687–692 (2005).
- Soss, S. E., Klevit, R. E. & Chazin, W. J. Activation of UbcH5c–Ub is the result of a shift in interdomain motions of the conjugate bound to U-box E3 ligase E4B. *Biochemistry* **52**, 2991–2999 (2013).
- Williamson, A. et al. Regulation of ubiquitin chain initiation to control the timing of substrate degradation. *Mol. Cell* **42**, 744–757 (2011).
- He, J. et al. Insights into degron recognition by APC/C coactivators from the structure of an Acm1–Cdh1 complex. *Mol. Cell* **50**, 649–660 (2013).
- Izawa, D. & Pines, J. The mitotic checkpoint complex binds a second CDC20 to inhibit active APC/C. *Nature* **517**, 631–634 (2015).
- Zeng, X. et al. Pharmacologic inhibition of the anaphase-promoting complex induces a spindle checkpoint-dependent mitotic arrest in the absence of spindle damage. *Cancer Cell* **18**, 382–395 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was funded by a Cancer Research UK grant to D.B. We thank W. J. Chazin and members of the Barford group for discussions, and X. Bai and S. Scheres for their help with RELION; C. Sava and S. Chen for EM facilities; P. Emsley for help with COOT; G. Murshudov for help with REFMAC; G. McMullan for assistance in movie data capture; J. Grimmett and T. Darling for computing; and A. Boland for advice with COOT and PyMol.

**Author Contributions** L.C. prepared grids, collected and analysed EM data and determined the three-dimensional reconstructions, fitted coordinates and built models, prepared figures and co-wrote the paper. Z.Z. designed and made constructs, performed biochemical analysis and purified proteins. J.Y. prepared and purified the complexes and performed biochemical analysis. S.H.McL. performed and analysed surface plasmon resonance experiments. D.B. directed the project, built models and co-wrote the paper.

**Author Information** EM maps have been deposited in the Electron Microscopy Data Bank under accession codes 2924 (APC/C<sup>Cdh1.Emi1</sup>), 2925 (APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup>) and 2926 (APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup>). APC/C<sup>Cdh1.Emi1</sup> coordinates have been deposited in the Protein Data Bank under accession number 4UI9. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.B. ([dbarford@mrc-lmb.cam.ac.uk](mailto:dbarford@mrc-lmb.cam.ac.uk)).



## METHODS

No statistical methods were used to predetermine sample size.

**Cloning, expression and purification of recombinant human APC/C.** The genes for recombinant human APC/C<sup>Cdh1.Emi1</sup> were cloned into a modified MultiBac system as described previously<sup>42</sup>. Two pU2 plasmids were prepared: pU2-1 harbouring genes for Apc1 and Apc11, and pU2-2 containing genes for Apc5, Apc8, Apc10, Apc13 and Apc15. Two pF2 plasmids were prepared: pF2-1 harbouring genes for Apc3, Apc6, Apc7, Apc12 and Apc16, and pF2-2 containing genes for Apc2, Apc4, Cdh1 and Emi1. A C-terminal *tev*-cleavable double Strep-II tag was added to Apc4 for purification of the complex. Two bacmids were prepared from these four plasmids, resulting in two viruses harbouring all 16 genes with each gene flanked by its own promoter and terminator. The APC/C<sup>Cdh1.Emi1</sup> ternary complex was co-expressed from these two viruses in High 5 insect cells. The complex was purified by a combination of Strep-Tactin (Qiagen), anion exchange chromatography Mono Q and Superose 6 size-exclusion chromatography (GE Healthcare) similar to ref. 42.

APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup> was cloned, expressed and purified as described previously<sup>4</sup>, except that Apc11 was replaced with Apc11 fused to UbcH10<sup>(C114K)</sup> at its C terminus by a 30-residue linker of GSA repeats. A similar version of APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup> fusion complex was also prepared using wild-type UbcH10 instead of UbcH10<sup>(C114K)</sup> for activity studies.

UbcH10<sup>LR</sup> was constructed by fusion of residues 154–222 of Ube2S to the C terminus of UbcH10<sup>(C114K)</sup> in the pET28 plasmid. An N-terminal His<sub>6</sub>-*tev* site was added to UbcH10<sup>LR</sup>. In the purified ubiquitin and UbcH10<sup>LR</sup> proteins the His<sub>6</sub> tag was retained on ubiquitin, but cleaved from UbcH10<sup>LR</sup>. The UbcH10<sup>LR</sup>-ubiquitin conjugate was generated as previously described<sup>4</sup> with minor modifications. UbcH10<sup>LR</sup> (200  $\mu$ M) was incubated with His<sub>6</sub>-tagged ubiquitin (400  $\mu$ M) at 25 °C for 16 h in a buffer containing 5  $\mu$ M UBA1, 3 mM ATP, 5 mM MgCl<sub>2</sub>, 0.8 mM TCEP, 150 mM NaCl, 50 mM Tris-HCl (pH 10.0). Additional 2 mM ATP and 2  $\mu$ M UBA1 were added to the reaction and incubated at 25 °C for another 8 h. The UbcH10<sup>LR</sup>-ubiquitin conjugate was separated from unconjugated UbcH10<sup>LR</sup> by Ni-NTA and separated from free ubiquitin by S75 size-exclusion chromatography (GE Healthcare). APC/C<sup>Cdh1.Hsl1</sup> was purified as described previously<sup>4</sup>. The purified APC/C<sup>Cdh1.Hsl1</sup> was concentrated to 3  $\mu$ M. APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup> was reconstituted by incubating APC/C<sup>Cdh1.Hsl1</sup> with UbcH10<sup>LR</sup>-ubiquitin conjugate in a 1:3 molar ratio on ice for 30 min in buffer of 20 mM HEPES (pH 8.0), 150 mM NaCl, 1 mM DTT. Just before applying to EM grids, the sample was cross-linked with 0.03% glutaraldehyde, and incubated on ice for 40 min. The reaction was quenched by adding 40 mM Tris-HCl (pH 8.0). The quenched sample was incubated on ice for 30 min before further purification with size-exclusion Superose 6. A similar UbcH10<sup>LR</sup> protein was made by using wild-type UbcH10 for use in activity tests. Surface plasmon resonance experiments for UbcH10<sup>LR</sup> and APC/C were performed as described previously<sup>4</sup>. The Y91A and S123G mutants of UbcH10 and UbcH10LR were prepared as for wild-type UbcH10.

For ubiquitination reactions, APC/C and its variants with mutant subunits (Apc1 and Apc8) were cloned, expressed and purified as for apo APC/C<sup>42</sup> by replacing the target subunit with its mutant. These included APC/C with Apc1<sup>L1197A/L1227A/I1230A/H1231A</sup>, APC/C with Apc8<sup>N339A/E374A</sup>, APC/C with Apc8<sup>N339A/E374R</sup> and APC/C with Apc8<sup>E410R</sup>.

Human Cdh1 was cloned into pU1 with an N-terminal His<sub>6</sub> followed by a haemagglutinin tag. On the basis of this plasmid, variants of Cdh1 mutations were made. These were Cdh1<sup>L160A/L161A</sup>, Cdh1<sup>Y91A/L99A</sup>, Cdh1<sup>L127A/F128A</sup>, Cdh1<sup>L94A/L95A/V153A</sup>, Cdh1<sup>L160A/L161A/Y91A/L99A</sup>, Cdh1<sup>L160A/L161A/L127A/F128A</sup>, Cdh1<sup>Y91A/L99A/L127A/F128A</sup>, Cdh1<sup>L160A/L161A/Y91A/L99A/L127A/F128A</sup>, Cdh1<sup>R47A</sup>, Cdh1<sup>R52A</sup>, Cdh1<sup>R493A</sup>, Cdh1<sup>T121A/S151A/S163A</sup>, Cdh1<sup>S40A/S151A/S163A</sup>, Cdh1<sup>S40A/T121A/S163A</sup>, Cdh1<sup>S40A/T121A/S151A</sup>, Cdh1<sup>S40A/T121A/S151A/S163A</sup> and Cdh1<sup>T32A/S36A/S70A/S138A/S146A</sup>.

All Cdh1 and its mutations were expressed in High Five insect cells. To purify Cdh1, the cells were lysed in a buffer of 50 mM Tris-HCl (pH 7.3), 500 mM NaCl, 20 mM imidazole (Buffer A), with a protease inhibitor cocktail. After loading, the Ni-NTA column was washed with ten column volumes of buffer A, followed by a gradient wash to 200 mM imidazole in buffer A and further washing with ten column volumes of 200 mM imidazole in buffer A. The protein was eluted in buffer B of 300 mM imidazole (pH 7.3), 500 mM NaCl. Cdh1 was finally purified using S-200 size exclusion chromatography in 20 mM HEPES (pH 7.0), 200 mM NaCl, 1 mM DTT.

Human ubiquitin complementary DNA was cloned into a pET28 plasmid with an N-terminal His<sub>6</sub> tag and TEV cleavage site. On the basis of this plasmid, ubiquitin<sup>136A</sup> and ubiquitin<sup>144A</sup> mutants were prepared. Wild-type and mutant ubiquitin were expressed in B834<sup>trre2</sup> strain and purified by Ni-NTA and S75 size-exclusion chromatography (GE Healthcare).

**APC/C ubiquitination assays.** Human APC/C and its different mutant derivatives, UBA1, UbcH10 and its mutants, Ube2S, ubiquitin and its mutants, and yeast Hsl1<sup>667–872</sup> were buffer exchanged into 20 mM HEPES (pH 8.0), 150 mM NaCl, 1 mM DTT. Human Cdh1 and its mutants were in 20 mM HEPES (pH 7.0), 200 mM NaCl, 1 mM DTT. Ubiquitination assays were performed as described<sup>4</sup>. For the UbcH10 competition assays, UbcH10<sup>C114K</sup> and UbcH10<sup>C114K-LR</sup> at the concentrations 0.001, 0.01, 0.1, 1 and 10  $\mu$ M were used to compete with UbcH10 (0.1  $\mu$ M). Cdh1 was at 20 nM. To test the UbcH10-RING domain interface, the following UbcH10 and UbcH10<sup>LR</sup> mutants were prepared and tested: Y91A and S123G. Single time point assays were at 10 min.

**Cdh1 phosphorylation assays.** CDK2-dependent phosphorylation of Cdh1 and its mutants was performed in buffer containing 50 mM HEPES (pH 8.0), 10 mM MgCl<sub>2</sub>, 2 mM ATP, 10 mM NaF, 10 mM glycerol phosphate and 150 mM NaCl. The reactions were incubated at 23 °C for 30 min with 500 nM of Cdh1 and 50 nM of active phosphorylated CDK2-cyclin A3. The phosphorylation was terminated using the CDK1/2 inhibitor (Calbiochem). Non-CDK2-treated Cdh1 samples were set up in parallel with phosphorylating reactions by adding buffer instead of CDK2-cyclin A3. The activity assays were performed as described for ubiquitination assays.

**Peptide assays.** Peptides for D-box lysine distance analysis were synthesized by Designer BioScience. The peptides were modified by N-terminal biotinylation. Peptides used were the following: Pep-1K (EQRPTALGDISNSKASGS); Pep-4K (EQRPTALGDISNSAGSKASGS); Pep-7K (EQRPTALGDISNSAGSKASGS); Pep-10K (EQRPTALGDISNSAGSKASGS); Pep-13K (EQRPTALGDISNSAGSKASGS); Pep-18K (EQRPTALGDISNSAGSKASGS); GEM-K50 (EQRPTALGDISNSAGSLVGRENSLGLSKRRHR); and CycB-K64 (GLRPTALGDIGNRVSEQLQARMPMRKEARPSATGS). The peptides were dissolved at a concentration of 6.25 mM in 100% dimethylsulfoxide (DMSO). The ubiquitination assays were performed as described above except that 125  $\mu$ M of peptide instead of Hsl1 was used as substrate. Previous work showed that at 100  $\mu$ M, D-box-only peptides bind the D-box receptor of APC/C<sup>Cdh1</sup> (refs 43, 44). The reactions were separated in 4–12% NuPAGE Bis-Tris gels (Life Technology) by 60 V for 2 h, followed by 150 V for 30 min for better resolution of peptides and ubiquitination products. After blotting, the membranes were blocked in 10% non-fat dry milk in PBS overnight at 4 °C. After washing the blot four times with PBS–2% Tween 20 buffer, ubiquitinated peptide products were visualized by western blotting with PBS-buffered streptavidin-HRP (Sigma) with 1% bovine serum albumin.

**Electron microscopy.** Freshly purified recombinant human APC/C complexes (APC/C<sup>Cdh1.Emi1</sup>, APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup> and APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup>) were first visualized by negative-stain electron microscopy to check homogeneity. Complexes were then diluted to ~200  $\mu$ g ml<sup>-1</sup> and aliquots of 2  $\mu$ l were applied to Quantifoil R2/2 grids coated with a second layer of home-made thin carbon film (estimated to be ~50 Å thick), treated with a 9:1 argon:oxygen plasma for 20 s before use. The grids were incubated for 30 s at 4 °C and 100% humidity and then blotted for 6 s and plunged into liquid ethane using an FEI Vitrobot III. The grids were loaded into an FEI Tecnai Polara electron microscope at an accelerating voltage of 300 kV. Images were taken manually using low-dose mode at a nominal magnification of  $\times 78,000$  (yielding a pixel size of 1.36 Å per pixel) for APC/C<sup>Cdh1.Emi1</sup> and APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup>, and  $\times 59,000$  (yielding a pixel size of 1.77 Å per pixel) for APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup>. Images were recorded using an FEI Falcon II direct electron detector with a defocus range of ~2.5–4  $\mu$ m. Each image was exposed for 2 s at dose rate of 27 electrons per square ångström per second at a nominal magnification of  $\times 78,000$  and 16 electrons per square ångström per second at a nominal magnification of  $\times 59,000$ . Thirty-four movie frames were recorded for each image as previously described<sup>45</sup>.

To prepare a negative-stain reconstruction of APC/C<sup>Cdh1.Hsl1</sup>-UbcH10 (with excess wild-type UbcH10), we incubated freshly purified APC/C<sup>Cdh1.Hsl1</sup> (0.04  $\mu$ M) with a 1,500-fold excess of UbcH10 (60  $\mu$ M, 200-fold in excess of dissociation constant  $K_d$ ) for 1 h on ice. Images were collected on an FEI Spirit electron microscope at an accelerating voltage of 120 kV at –1.5  $\mu$ m defocus. The map was reconstructed using RELION without contrast transfer function correction.

**Image processing.** All movie frames were first aligned by the motioncorr program<sup>46</sup> before subsequent processing. Contrast transfer function parameters were calculated with CTFFIND3 (ref. 47). Particles in 264 pixels  $\times$  264 pixels (or 200 pixels  $\times$  200 pixels for data collected at  $\times 59,000$  magnification) were initially selected automatically by e2boxer.py in EMAN2 (ref. 48) from a small portion of the same data set (typically ~300 micrographs). After two-dimensional classification using RELION<sup>49</sup>, ten images in different views were selected from the two-dimensional averages and used as reference for automatic particle picking for the whole data set by RELION 1.3 (ref. 50). The following procedures were performed in RELION to exclude bad particles from the final

reconstruction. First, particles in each micrograph were displayed and manually screened to delete wrongly picked ice contamination. Second, extracted particles were sorted by similarity to reference images, and particles with low z-scores were deleted (roughly 10% particles were deleted in this step). Third, two-dimensional classification was performed and particles in bad classes (with poor structural features) were removed (roughly 10–20% of particles were deleted in this step). The resultant particles were then used for three-dimensional classifications. Conformational heterogeneity and remaining bad particles were excluded from the final reconstruction in this step. A summary of particles used for final refinements is listed in Extended Data Table 1c.

Beam-induced particle motion was corrected using statistical movie processing in RELION as previously described<sup>45</sup>. For APC/C, a running average of nine movie frames for frames within the first 32-electron dose and a standard deviation of one pixel for the translational alignment were used. This step improved the resolution from 4.2 Å to 3.9 Å. Following this, particle polishing in RELION 1.3 was used to improve the resolution further<sup>51</sup>. This program fits all movie frames as linear tracks, considering movements of neighbouring particles to improve signal for smaller particles. Furthermore, each frame is weighted by a B-factor (estimated by each frame reconstruction) to treat radiation damage problems. Particle polishing was then used to improve the resolution from 3.9 Å to 3.6 Å. All resolution estimations were from gold-standard Fourier shell correlation (FSC) calculations to avoid over-fitting and reported resolutions are based on the FSC = 0.143 criterion<sup>52,53</sup>. Final FSC curves were calculated using a soft spherical mask (with a five-pixel fall-off) of the two independent reconstructions. To visualize high-resolution details, all density maps were corrected for modulation transfer function of the detector and sharpened by applying negative B-factors, estimated using automated procedures<sup>52</sup>.

Motion correction was performed after three-dimensional classification for APC/C<sup>Cdh1,Emi1</sup>, which showed ~60% ternary complex and ~40% apo state APC/C. For APC/C<sup>Cdh1,Hsl1,UbcH10-UB</sup> and APC/C<sup>Cdh1,Hsl1,Apc11-UbcH10</sup>, all particles were refined against one model and motion corrected as described above. The polished particles were used for three-dimensional classification as shown in Extended Data Figs 6 and 7. The first cycle of classification was done by global search and a sampling angular interval of 7.5° to sort out large conformational changes; cycles 2–4 were executed by local search within 15° and a smaller sampling angular interval of 3.7°. Three-dimensional classifications were run for 25 iterations with class number set at 10 and regularization parameter T set at 4.

To further improve the density map of flexible parts (RING domain and UbcH10) in the complexes, a soft mask was used for local alignment during refinement. A soft mask was generated according to combined coordinates of Apc1/Apc2/Apc4/Apc5/Apc11/UbcH10 (approximately half of the mass) and a restrained angular search of 3° was used. Local resolution was estimated using ResMap<sup>54</sup>.

**Model building of APC/C.** Modelling of the structures of APC/C was performed in COOT<sup>55</sup>. Initially, available crystal structures and homology models as described in our previous work<sup>4</sup> were fitted into the cryo-EM map in Chimera. These structures included (1) crystal structures of human Apc10 (residues 2–162, Protein Data Bank (PDB) accession number 1JHJ)<sup>56</sup>, (2) N-terminal seven helices of Apc7 (residues 21–166, PDB accession number 3FFL)<sup>57</sup>, (3) *Schizosaccharomyces pombe* Cut9–Hcn26 complex (Apc6–Apc12 homologue, PDB accession number 1XPI)<sup>58</sup>, (4) N-terminal dimerization domains of *Encephalitozoon cuniculi* Cdc27 (Apc3 homologue, PDB accession number 3KAE)<sup>59</sup>, (5) *S. pombe* Cut23 (Apc8 homologue, PDB accession number 3ZN3)<sup>60</sup>, (6) Cdh1-D-KEN (PDB accession number 4BH6)<sup>39</sup>, (7) Apc5 TPR domain and C-terminal TPR helices of Apc3, Apc8 and Apc7 (model using the crystal structure of *S. pombe* Cut9 as template), (8) crystal structures of Apc4 and Apc5<sup>NTD</sup> (residues 1–160) (unpublished data) and (9) NMR structure of Emi1 ZBR domain<sup>13</sup> (PDB accession number 2M6N) and Apc11 RING domain<sup>26</sup> (PDB accession number 4R2Y). For Apc2–Apc11<sup>N</sup>, the structure was modelled using a crystal structure of Cul1 (PDB accession number 1U6G)<sup>61</sup> as a template. For Apc1 PC repeats, the structure was predicted with I-TASSER<sup>62</sup> using the PC domain structure of Rpn2 (PDB accession number 4ADY)<sup>63</sup> as a reference. All fitted structures were rebuilt according to the cryo-EM map. The majority of Apc1, small subunits (Apc13, Apc15, Apc16), IR C-terminal tails of Apc10 and Cdh1, N-terminal domain of Cdh1 (Cdh1<sup>NTD</sup>) and some loops not seen in crystal structures and homology were built *ab initio*. Details are given in Extended Data Table 2.

The model was refined by REFMAC version 5.8 (ref. 64). A REFMAC weight of 0.04 was defined by cross-validation using half reconstructions<sup>65</sup>. The resolution limit for refinement was set at 3.5 Å. All high-resolution crystal structures or

NMR structures were used for secondary structure restraints. A summary of model refinement is listed in Extended Data Table 1a.

Rigid-body fitting was performed to UbcH10 by the ‘fit in map’ program in Chimera<sup>66</sup>. The crystal structure of the Rnf4–UbcH5a–ubiquitin heterotrimeric complex (PDB accession number 4AP4)<sup>29</sup> was used to model ubiquitin not seen directly from the reconstruction owing to flexibility. A summary of model building is listed in Extended Data Table 2.

All regions of the APC/C EM density map were assigned to APC/C subunits. However, the intrinsically disordered regions of some APC/C subunits were not visible in the EM map (defined in Extended Data Table 2). Altogether, 78% APC/C residues were fitted to the EM map. Inter-subunit contact areas were calculated using CCP4 (ref. 67).

**Map visualization.** Figures and movies were generated using Pymol and Chimera<sup>66</sup>.

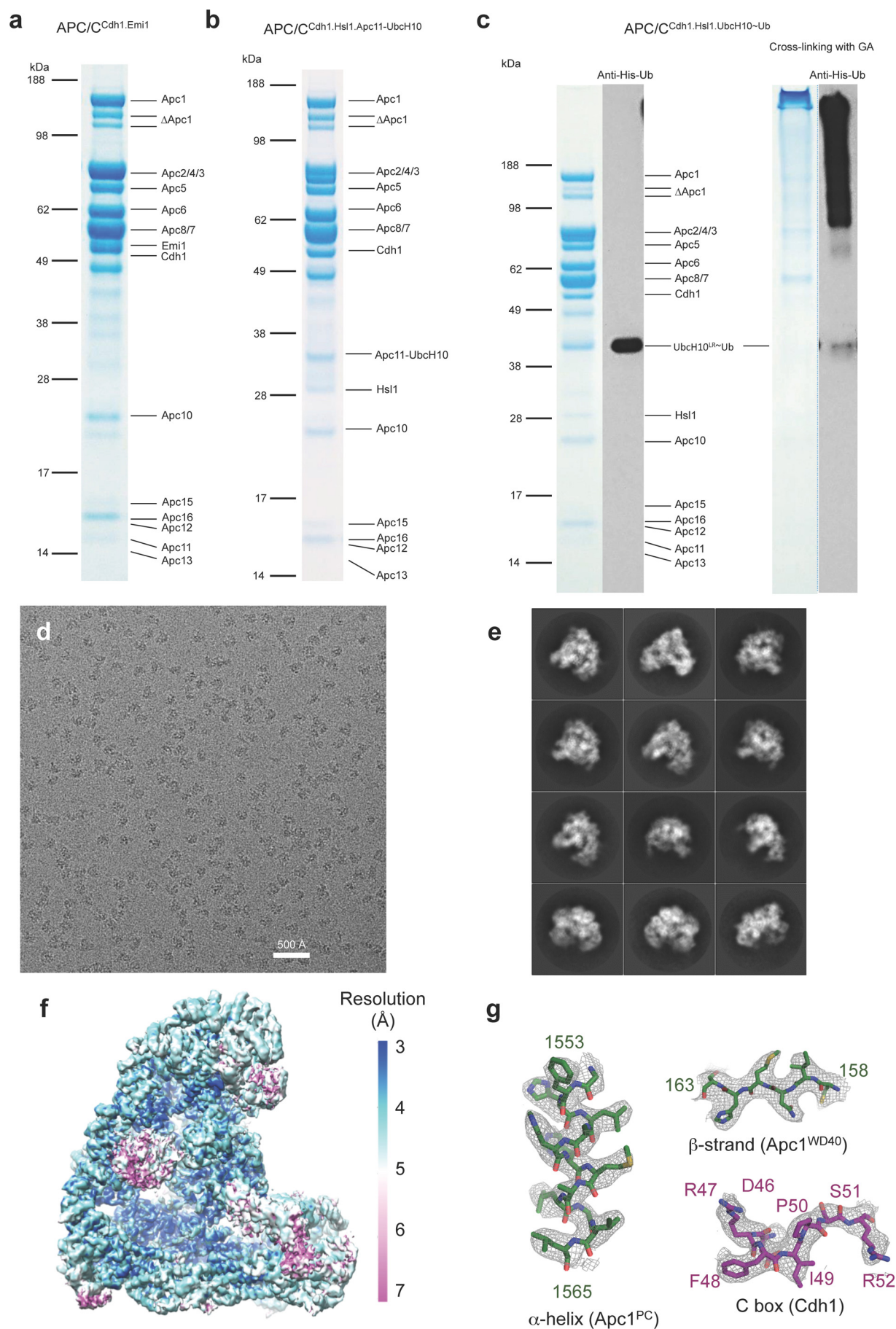
**Structural conservation.** Structural conservation was calculated using CONSURF<sup>68</sup> on the basis of a multiple sequence alignment of *Homo sapiens*, *Mus musculus*, *Xenopus laevis*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *S. cerevisiae*, *S. pombe*, APC/C subunit sequences.

**Protein similarity searches.** Protein similarity searches were performed using DALI<sup>69</sup>, and protein alignment figures created using ALSCRIPT<sup>70</sup>.

42. Zhang, Z. *et al.* Recombinant expression, reconstitution and structure of human anaphase-promoting complex (APC/C). *Biochem. J.* **449**, 365–371 (2013).
43. Kraft, C., Vodermaier, H. C., Maurer-Stroh, S., Eisenhaber, F. & Peters, J. M. The WD40 propeller domain of Cdh1 functions as a destruction box receptor for APC/C substrates. *Mol. Cell* **18**, 543–553 (2005).
44. da Fonseca, P. C. *et al.* Structures of APC/C(Cdh1) with substrates identify Cdh1 and Apc10 as the D-box co-receptor. *Nature* **470**, 274–278 (2011).
45. Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife* **2**, e00461 (2013).
46. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods* **10**, 584–590 (2013).
47. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
48. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
49. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
50. Scheres, S. H. Semi-automated selection of cryo-EM particles in RELION-1.3. *J. Struct. Biol.* **189**, 114–122 (2015).
51. Scheres, S. H. Beam-induced motion correction for sub-megadalton cryo-EM particles. *eLife* **3**, e03665 (2014).
52. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
53. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nature Methods* **9**, 853–854 (2012).
54. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nature Methods* **11**, 63–65 (2014).
55. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
56. Wendt, K. S. *et al.* Crystal structure of the APC10/DOC1 subunit of the human anaphase-promoting complex. *Nature Struct. Biol.* **8**, 784–788 (2001).
57. Han, D. *et al.* Crystal structure of the N-terminal domain of anaphase-promoting complex subunit 7. *J. Biol. Chem.* **284**, 15137–15146 (2009).
58. Zhang, Z., Kulkarni, K., Hanrahan, S. J., Thompson, A. J. & Barford, D. The APC/C subunit Cdc16/Cut9 is a contiguous tetratricopeptide repeat superhelix with a homo-dimer interface similar to Cdc27. *EMBO J.* **29**, 3733–3744 (2010).
59. Zhang, Z. *et al.* Molecular structure of the N-terminal domain of the APC/C subunit Cdc27 reveals a homo-dimeric tetratricopeptide repeat architecture. *J. Mol. Biol.* **397**, 1316–1328 (2010).
60. Zhang, Z. *et al.* The four canonical tpr subunits of human APC/C form related homo-dimeric structures and stack in parallel to form a TPR suprahelix. *J. Mol. Biol.* **425**, 4236–4248 (2013).
61. Zheng, N. *et al.* Structure of the Cul1–Rbx1–Skp1–F box<sup>Skp2</sup> SCF ubiquitin ligase complex. *Nature* **416**, 703–709 (2002).
62. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**, 725–738 (2010).
63. He, J. *et al.* The structure of the 26S proteasome subunit Rpn2 reveals its PC repeat domain as a closed toroid of two concentric alpha-helical rings. *Structure* **20**, 513–521 (2012).
64. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
65. Fernandez, I. S., Bai, X. C., Murshudov, G., Scheres, S. H. & Ramakrishnan, V. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* **157**, 823–831 (2014).
66. Yang, Z. *et al.* UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J. Struct. Biol.* **179**, 269–278 (2012).
67. Collaborative. Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).

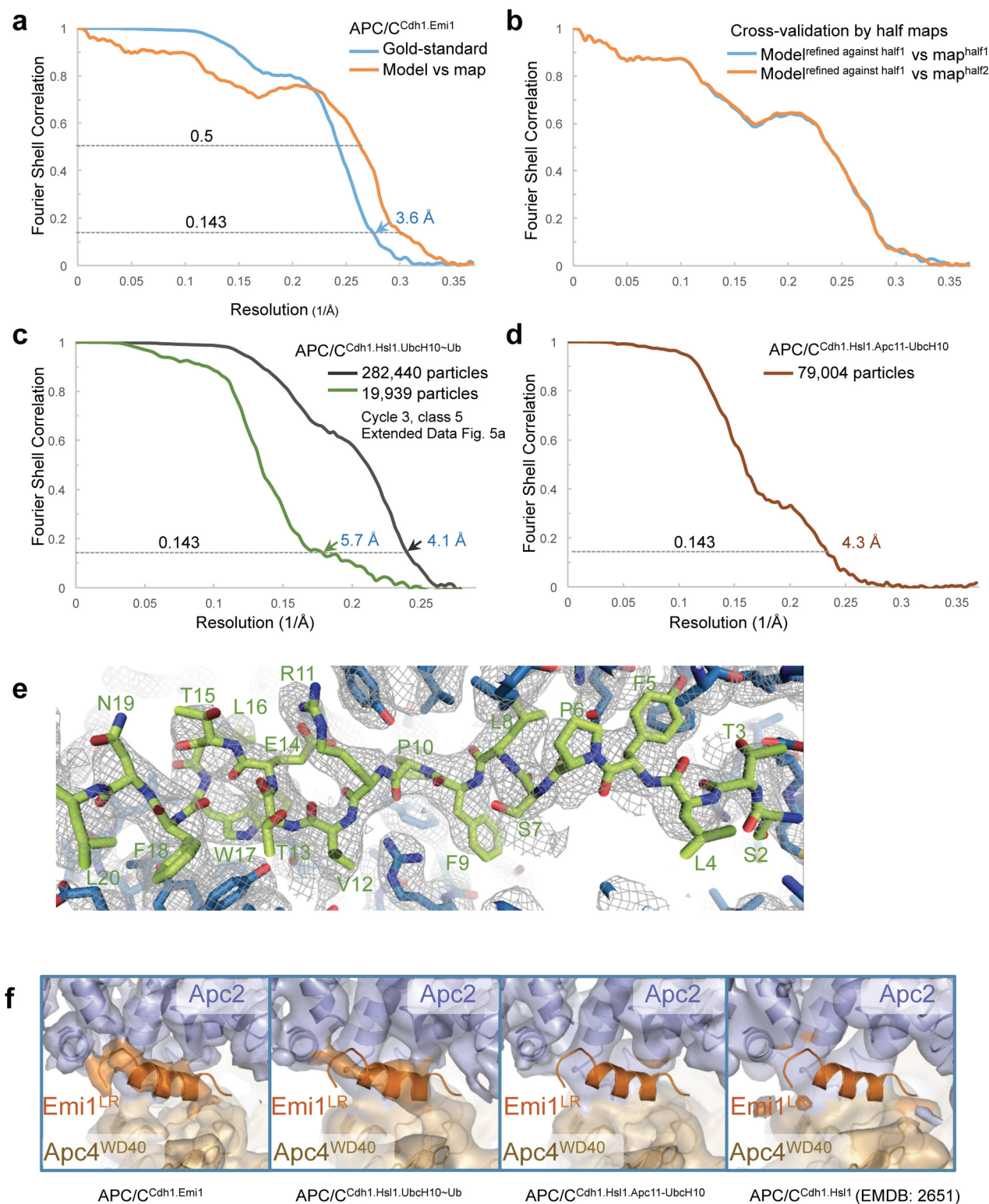
68. Landau, M. *et al.* ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **33**, W299–W302 (2005).
69. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780–2781 (2008).
70. Barton, G. J. ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40 (1993).
71. Hegemann, B. *et al.* Systematic phosphorylation analysis of human mitotic protein complexes. *Sci. Signal.* **4**, rs12 (2011).
72. Kraft, C. *et al.* Mitotic regulation of the human anaphase-promoting complex by phosphorylation. *EMBO J.* **22**, 6598–6609 (2003).
73. Steen, J. A. *et al.* Different phosphorylation states of the anaphase promoting complex in response to antimitotic drugs: a quantitative proteomic analysis. *Proc. Natl Acad. Sci. USA* **105**, 6069–6074 (2008).





**Extended Data Figure 1 | Preparations and EM images of APC/C complexes.** **a**, Coomassie-blue-stained SDS gel of APC/C<sup>Cdh1.Emi1</sup>. **b**, Coomassie-blue-stained SDS gel of APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup>. **c**, Coomassie-blue-stained SDS gel and western blot analysis (anti-His antibody—only ubiquitin in the complex contains His tag) of APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup> without or with cross-linking by glutaraldehyde. **d**, A typical cryo-EM micrograph of APC/C<sup>Cdh1.Emi1</sup>;

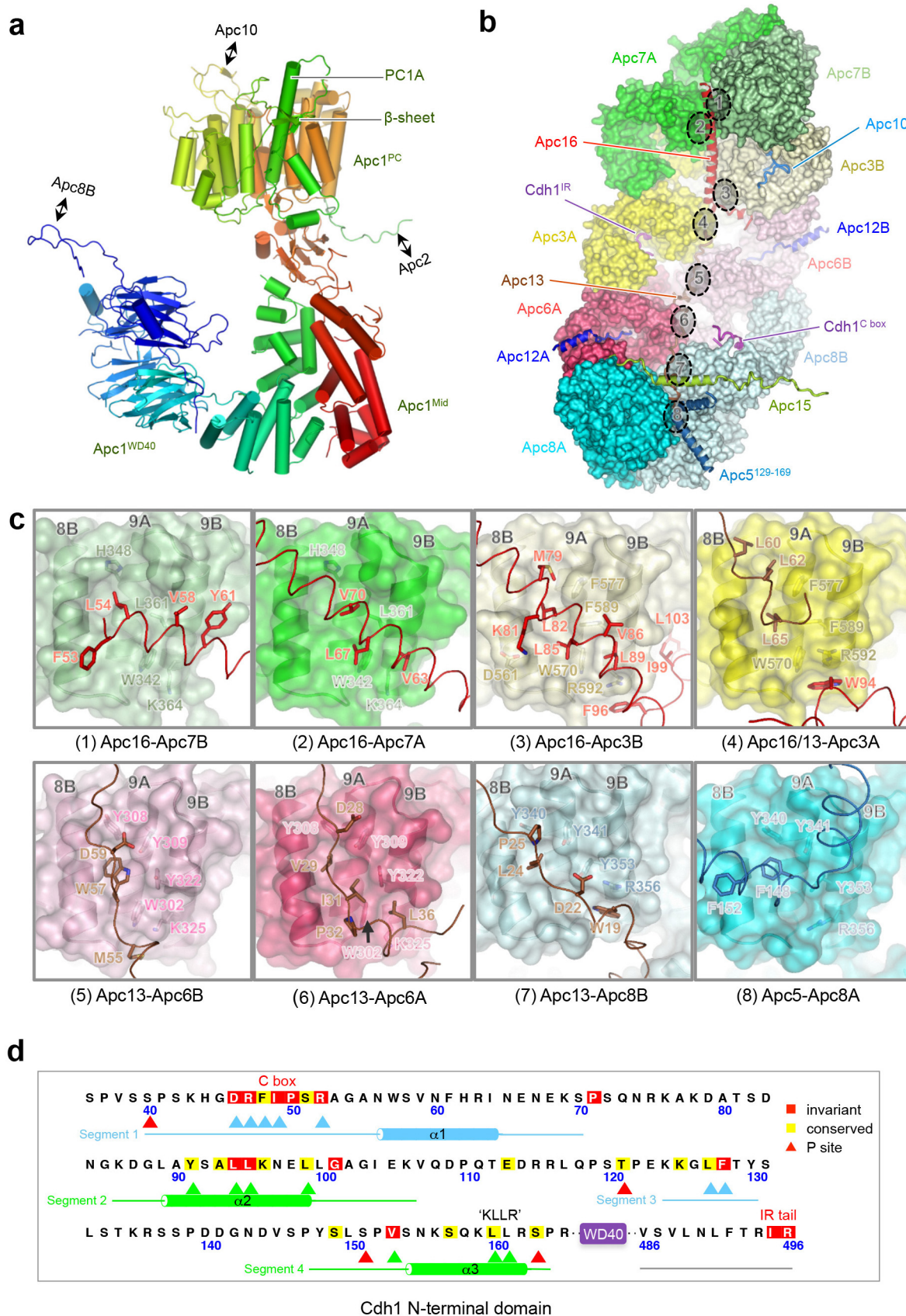
representative of 3,328 micrographs. **e**, Gallery of two-dimensional averages of APC/C<sup>Cdh1.Emi1</sup> showing different views; representative of 100 two-dimensional averages. **f**, Local resolution map of APC/C<sup>Cdh1.Emi1</sup> showing resolution range. **g**, Details of EM density for segments of  $\alpha$ -helix and  $\beta$ -strand of Apc1 and the C box of Cdh1.



**Extended Data Figure 2 | Resolution estimation and example of *de novo* model building.** **a**, Gold-standard FSC curve and FSC curve between cryo-EM map and final atomic model of the APC/C<sup>Cdh1.Emi1</sup>. **b**, Cross-validation of model refinement by half maps. Shown are FSC curves between the atomic model and the half map (map<sup>half1</sup>) it was refined against, and FSC curves between the atomic model and the other half map (map<sup>half2</sup>) that was not used during refinement. **c**, Gold standard FSC curve of APC/C<sup>Cdh1.Hsl1.UbcH10~Ub</sup>.

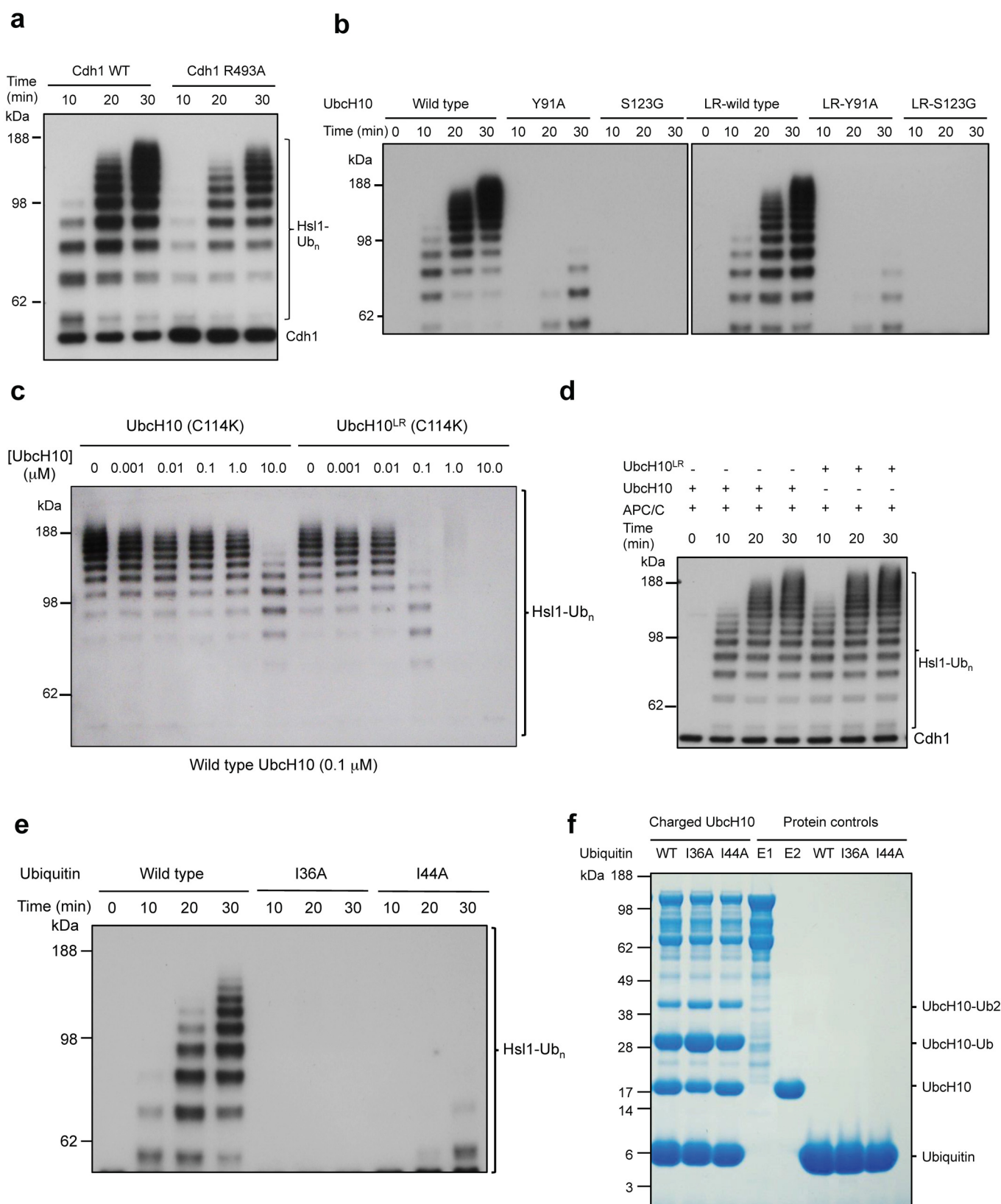
**d**, Gold-standard FSC curves of human APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup>. **e**, *De novo* model building of Apc15 N-terminal loop. The surrounding Apc5 residues are also shown. **f**, EM density for the LR tail common to Emi1 and Ube2S is only observed in APC/C complexes with subunits incorporating the LR tail. (1) APC/C<sup>Cdh1.Emi1</sup>, (2) An APC/C complex with an LR tail-bearing subunit (UbcH10<sup>LR</sup> of APC/C<sup>Cdh1.Hsl1.UbcH10~Ub</sup>), (3) No LR tail density in APC/C<sup>Cdh1.Hsl1.Apc11-UbcH10</sup> fusion and (4) APC/C<sup>Cdh1.Hsl1</sup> (ref. 4).





**Extended Data Figure 3 | Apc1 structure and the TPR lobe interacts with multiple subunits.** **a**, Cartoon of Apc1 colour-ramped from blue to red for the N to C termini. Insertions that interact with Apc2, Apc8 and Apc10 are labelled. Apc1<sup>Mid</sup> adopts a novel architecture. **b**, TPR lobe with TPR subunits shown as surface representations, with the small TPR-accessory subunits (Apc12, Apc13, Apc15 and Apc16), a segment of Apc5, IR tails of Cdh1 and Apc10, and the Cdh1 C box that interact with the TPR lobe are shown as cartoons. The N termini of Apc12, Apc13 and Apc15 are buried. The eight structurally homologous and symmetry-equivalent sites on the TPR lobe that

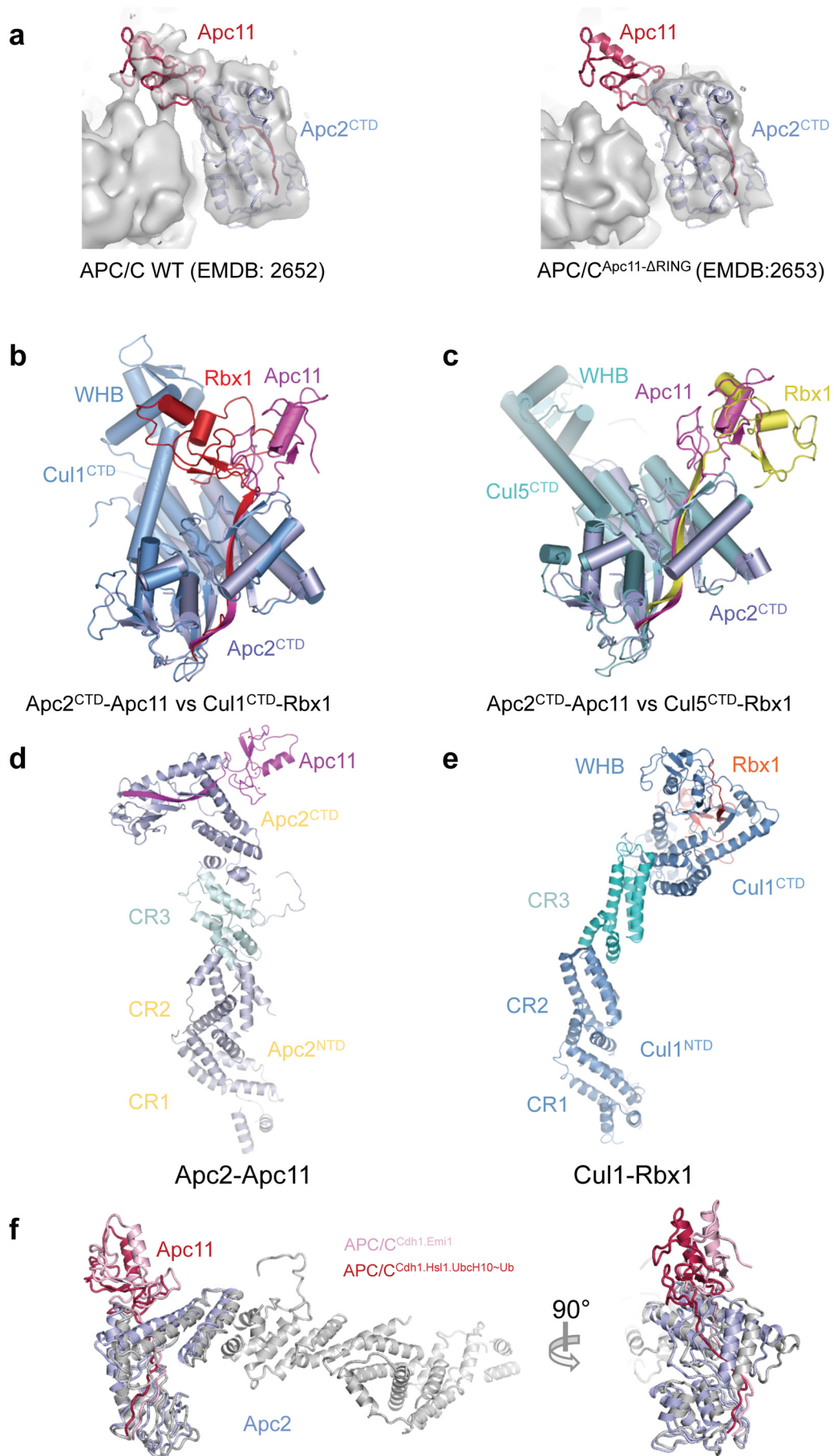
bind Apc13, Apc16 and Apc5 are indicated and shown in detail in **c**. The view is similar to Fig. 1a. **c**, The eight TPR subunits interact with Apc13, Apc16 and Apc5 mainly through contacts to four conserved aromatic residues present on most TPR subunits (Y308, Y309, W302, Y322 of Apc6 (panels 5 and 6)). **d**, Sequence of the ordered region of Cdh1<sup>NTD</sup> bound to Apc1 and Apc8 (ordered regions shown as lines and  $\alpha$ -helices). Critical Apc1 and Apc8 contact residues are indicated with green and blue arrows. Phosphorylation sites are indicated with red arrows.



**Extended Data Figure 4 | APC/C ubiquitination assays.** **a**, Mutation of Arg493 of the IR tail reduces APC/C<sup>Cdh1</sup> activity. **b**, Mutations at the RING domain interface of UbchH10 and UbchH10<sup>LR</sup> disrupt ubiquitination activity. **c**, Ubiquitination assay shows that both UbchH10(C114K) and UbchH10<sup>LR</sup>(C114K) compete with wild-type UbchH10. UbchH10<sup>LR</sup>(C114K) is a more potent inhibitor. **d**, The APC/C–UbchH10-mediated substrate

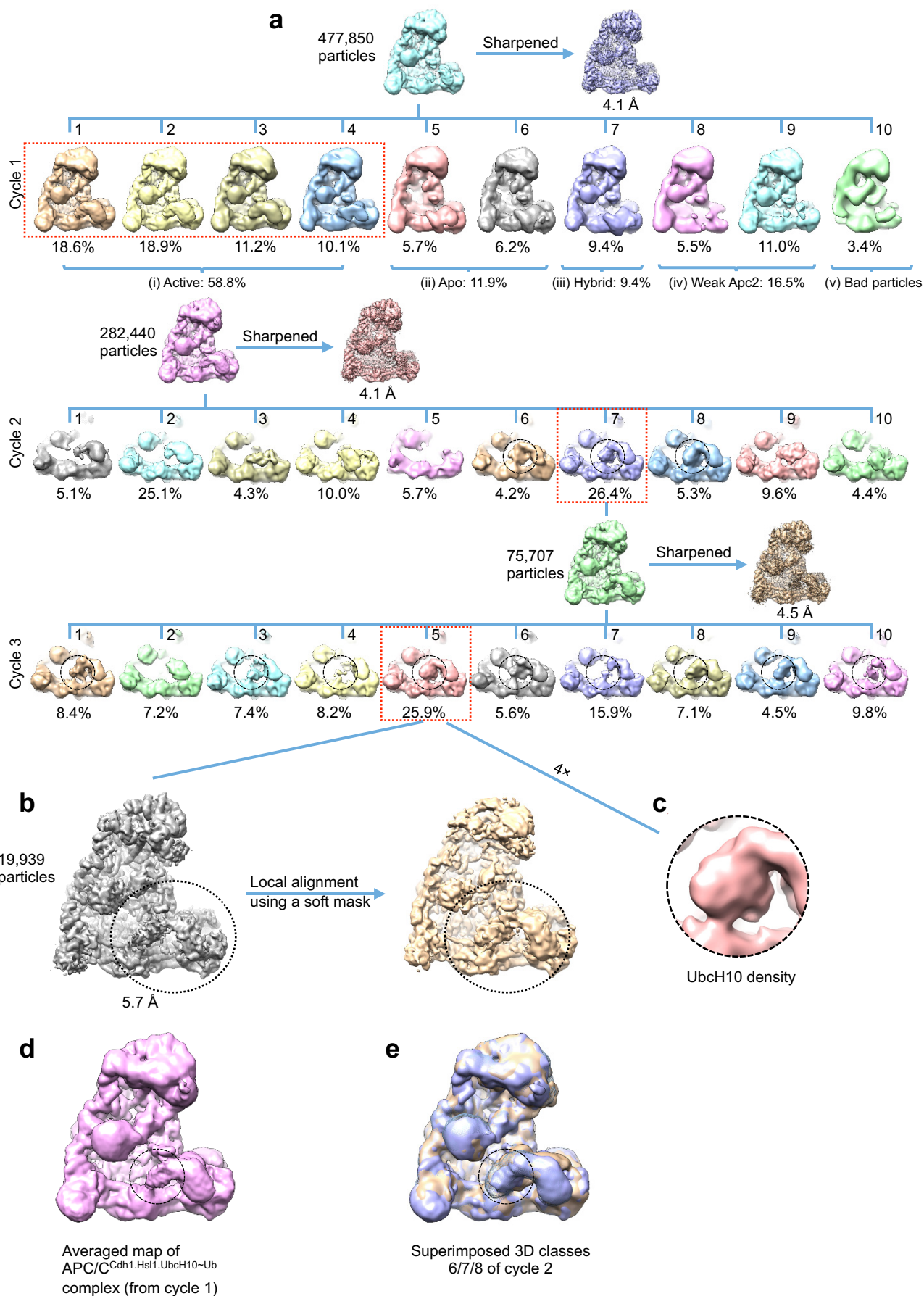
ubiquitination activities of UbchH10 and UbchH10<sup>LR</sup> are indistinguishable. **e**, The ubiquitin (I36A) and ubiquitin (I44A) mutants were defective for APC/C–UbchH10-mediated substrate ubiquitination. Experiments in Extended Data Fig. 4a–e were replicated three times. **f**, UbchH10 charging by the ubiquitin (I36A) and ubiquitin (I44A) mutants was unchanged relative to wild-type ubiquitin.





**Extended Data Figure 5 | The position of Apc11<sup>RING</sup> in the APC/C is more similar to Rbx1<sup>RING</sup> of activated cullin-Rbx1 structures.** **a**, Identification of Apc11 in apo APC/C. Left panel: EM density map for apo APC/C with the coordinates of Apc2<sup>CTD</sup>-Apc11 fitted (from APC/C<sup>Cdh1.Emi1</sup> structure). Right panel: EM density for APC/C<sup>Apc11-ΔRING</sup>. The difference density corresponds to Apc11<sup>RING</sup>. EM density maps from ref. 4. **b**, Superimposed Apc2<sup>CTD</sup> onto Cul1<sup>CTD</sup> (PDB accession number 1LDK)<sup>61</sup>. **c**, Superimposed Apc2<sup>CTD</sup> onto Cul5<sup>CTD</sup> (PDB accession number 3DQV)<sup>27</sup>. In the inactive conformation of Cul1-Rbx1, Rbx1<sup>RING</sup> packs against WHB. In APC/C<sup>Cdh1.Emi1</sup> the location of Apc11<sup>RING</sup> remains in contact with Apc2<sup>CTD</sup> but has rotated ~180° relative to

inactive CRL structures, being similar to the swung out conformation of Rbx1<sup>RING</sup> of neddylylated and activated Cul5-Rbx1 (ref. 27). **d**, **e**, The relative orientation of Apc2<sup>NTD</sup> and Apc2<sup>CTD</sup> is also dramatically different from Cul1 (ref. 61). This is due to a 70° rotation within cullin repeat 3 (between helices A-B and C-D-E), and a ~20° rotation around the 4HB-cullin repeat 3 interface. Similar less pronounced structural variations are observed within the CRL family. **d**, Apc2-Apc11 (this study). **e**, Cul1-Rbx1 (PDB accession number 1LDK)<sup>61</sup>. **f**, The position of the Apc2<sup>CTD</sup>-Apc11 module differs slightly about the Apc2<sup>NTD</sup>-Apc2<sup>CTD</sup> interface between APC/C<sup>Cdh1.Emi1</sup> and APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup>.

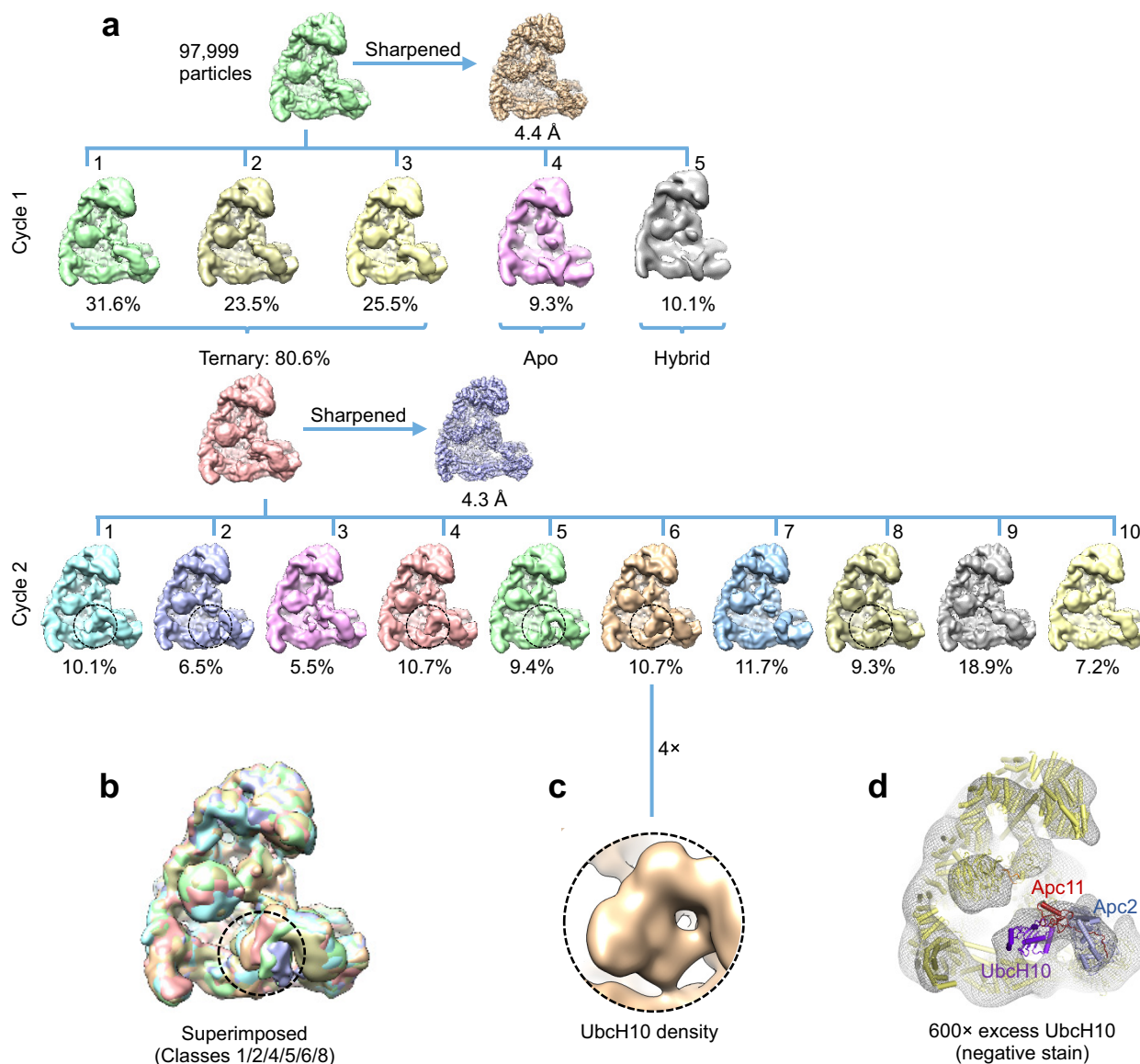




### Extended Data Figure 6 | Three-dimensional classification of APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup>

**a**, The three-dimensional classification (cycle 1) started with 477,850 motion-corrected particles, which were divided into ten classes. The resultant classes were grouped into five categories: (1) 58.8% in the active ternary state with coactivator and substrate (Hsl1); (2) 11.9% in the apo-inactive state; (3) 9.4% in a class with Cdh1 bound but with the catalytic module in the apo-inactive conformation (hybrid state); (4) 16.5% with weak Apc2 density; and (5) 3.4% were a poor reconstruction due to bad particles. Examination of the hybrid state (3) showed that density for Apc1<sup>WD40</sup> was absent, explaining the lack of Cdh1-induced conformational change of the catalytic module. Particles in the ternary state (reconstructed to an overall resolution of 4.1 Å) were subjected to further three-dimensional classification (cycle 2). A major class (class 7, 26.4% of particles) showed improved density

for UbcH10 (circled), and cycle 3 classification was performed on particles in this class. The major class of cycle 3 (class 5, 25.9% particles) showed further improved UbcH10 density. Further three-dimensional classification of this class did not improve the UbcH10 density. **b**, Particles in the best class (cycle 3, class 5, 19,939 particles) were refined in RELION and resulted in a map at 5.7 Å resolution (Extended Data Figure 2c). The UbcH10 density was improved by local alignment using a soft mask (indicated by circles) as described in Methods. **c**, Enlarged view of UbcH10 density. **d**, Enlarged view of the averaged APC/C<sup>Cdh1.Hsl1.UbcH10-Ub</sup> reconstruction from cycle 1 of the three-dimensional classification (59% of particles). UbcH10 density is circled. **e**, Superimposition of classes 6, 7 and 8 of cycle 2 of the three-dimensional classification (from **a**), showing the structural variability of the three three-dimensional classes that indicate UbcH10 density. UbcH10 density is circled.



**Extended Data Figure 7 | Three-dimensional classification of APC/C<sup>Cdh1.Hsl1.Apc11</sup>-Ubch10**

**a**, The three-dimensional classification started with 97,999 motion-corrected particles, which were divided into five classes. The resultant classes were grouped into three categories: (1) 80.6% in the active ternary state with coactivator and substrate (Hsl1); (2) 9.3% in the apo state; and (3) 10.1% in a hybrid state. Particles in the active ternary state (reconstructed to an overall resolution of 4.3 Å) were subjected to cycle 2 classification with ten classes. Ubch10 density in the resultant classes is

indicated with circles. **b**, Classes with Ubch10 density in cycle 2 classification are superimposed, showing variability of Ubch10. **c**, Enlarged view of Ubch10 density. **d**, A negative-stain EM reconstruction of an APC/C<sup>Cdh1.Hsl1</sup> complex at ~25 Å resolution with a 1,500-fold excess of Ubch10. The molecular surface is shown as a mesh representation and the coordinates of the APC/C<sup>Cdh1.Hsl1.Ubch10-Ub</sup> were docked into the EM reconstruction. The Ubch10 coordinates fit new EM density proximal to Apc11.

Extended Data Table 1 | EM data collection and processing statistics, and structure refinement statistics

**a. Statistics of APC/C<sup>Cdh1.Emi1</sup> structure determination****Data collection**

|                     |                    |
|---------------------|--------------------|
| EM:                 | FEI Polara, 300keV |
| Detector:           | FEI Falcon II      |
| Pixel size (Å):     | 1.36               |
| Defocus range (μm): | 1.5–4              |

**Reconstruction**

|                                    |            |
|------------------------------------|------------|
| Software:                          | RELION 1.3 |
| Accuracy of rotations (degrees):   | 1.073      |
| Accuracy of translations (pixels): | 0.691      |
| Final Resolution (Å):              | 3.6        |

**Refinement**

|                                   |            |
|-----------------------------------|------------|
| Software:                         | RefMac 5.8 |
| Refmac weight:                    | 0.04       |
| Resolution limits (Å):            | 3.5        |
| Residue number                    | 8408       |
| Average Fourier shell correlation | 0.8315     |
| R factor:                         | 0.222      |
| Rms BondLength:                   | 0.0135     |
| Rms BondAngle:                    | 1.8098     |

**Validation**

|                    |               |
|--------------------|---------------|
| Ramachandran plot: |               |
| Preferred:         | 7669 (92.88%) |
| Allowed:           | 352 (4.26%)   |
| Outliers:          | 236 (2.86%)   |

**b. Statistics of all cryo-EM data**

| Samples  | Micro-graphs | Particles (after movie correction)            |   | Resolution (Å) |
|--|--------------|---|---|----------------|
| APC/C <sup>Cdh1.Emi1</sup>                               | 3328         | In total: 363,774                             |   | -              |
|  |              | Ternary                                       | 202,084 (55.6%)                                   | 3.6            |
|  |              | Apo   | 161,690 (44.4%)                                   | 4.3            |
|  |              | *Others (Hybrid, Weak Apc2 and Bad particles) | *52,178 particles deleted before movie correction | -              |
| APC/C <sup>Cdh1.Hsl1</sup> - UbcH10 <sup>LR</sup> ~Ub    | 4524         | In total: 477,850                             |   | -              |
|  |              | Ternary                                       | 282,440 (59.1%)                                   | 4.1            |
|  |              | Apo   | 56,865 (11.9%)                                    | -              |
|  |              | Others (Hybrid, Weak Apc2 and Bad particles)  | 138,545 (29.0%)                                   | -              |
| APC/C <sup>Cdh1.Hsl1</sup> - UbcH10 <sup>Apc11</sup> ~Ub | 869          | In total: 97,999                              |   | -              |
|  |              | Ternary                                       | 79,004 (80.6%)                                    | 4.3            |
|  |              | Apo   | 9,114 (9.3%)                                      | -              |
|  |              | Others (Hybrid, Weak Apc2 and Bad particles)  | 9,801 (10.1%)                                     | -              |



Extended Data Table 2 | Summary of model building of APC/C subunits

| Subunit          | Length | Domain/Region 1  | Domain/Region 2  | Domain/Region 3  | Disordered regions  | Phosphorylation sites   |
|------------------|--------|--|--|--|---|---|
| <b>Apc1</b>      | 1,944  | WD40 domain (1-612)<br>De novo                                       | Mid-N (613-986)<br>Mid-C (1617-1944)<br>De novo  | PC domain (1013-1616)<br>Homology. PDB: 4ADY<br>Inserts: De novo   | 1-10,59-70,135-145,193-205,227-235,296-401,415-423,514-581,671-756,815-838,883-923,987-1012,1334-1346,1438-1451,1712-1726,1930-1944 | 202,286,291,341,343,355,362,364,373,377,537,571,681,688                                     |
| <b>Apc2</b>      | 822    | NTD (1-432)<br>Homology. PDB: 1LDK                                   | CTD(433-822)<br>Homology. PDB: 1LDK  | -  | 1-14, 30-52,66-69,103-106, 192-195, 221-231,305-319,444-450,459-475<br>731-822 (WHB)  | 218,314,532,534   |
| <b>Apc3</b>      | 824    | TPR dimer interface<br>TPR motifs 1-7 (1-535)<br>Homology. PDB: 3KAE | TPR superhelix<br>TPR motifs 8-14 (536-824)<br>Homology. PDB: 2XPI                     | -  | 171-450,767-824 (A)<br>171-450,782-824 (B)  | 205,209,220,244,264,291,302,304,313,334,336,364,366,369,383,384,386,426,430,435,438,446,447 |
| <b>Apc4</b>      | 808    | WD40 domain/4HB<br>X ray. PDB: 5BPW<br>Inserts: De novo              | -  | -  | 429-438,458-469,758-808   | 469,777   |
| <b>Apc5</b>      | 755    | NTD (1-169)<br>Homology. PDB: 5BPZ                                   | TPR superhelix<br>TPR motifs 1-13 (206-755)<br>Homology. PDB: 2XPI<br>Inserts: De novo | -  | 1-26,170-205  | 15,232,195  |
| <b>Apc6</b>      | 620    | TPR dimer interface<br>TPR motifs 1-7 (1-261)<br>Homology. PDB: 2XPI | TPR superhelix<br>TPR motifs 8-14 (262-620)<br>Homology. PDB: 3HYM                     | -  | 96-126,528-620 (A)<br>96-123,534-620 (B)  | 112,490,559,560,580,581,585,586,599   |
| <b>Apc7</b>      | 599    | TPR dimer interface<br>TPR motifs 1-3 (21-166)<br>X ray. PDB: 3FFL   | TPR dimer interface<br>TPR motifs 4-7 (167-359)<br>Homology. PDB: 2XPI                 | TPR superhelix<br>TPR motifs 8-14 (360-599)<br>Homology. PDB: 2XPI | 1-35,111-131,553-599 (A)<br>1-35,111-131,541-599 (B)  | 51,57,64,119,120,125  |
| <b>Apc8</b>      | 597    | TPR dimer interface<br>TPR motifs 1-7 (1-287)<br>Homology. PDB: 3ZN3 | TPR superhelix<br>TPR motifs 8-14 (288-597)<br>Homology. PDB: 2XPI                     | -  | 1-25,501-508,558-597 (A)<br>1-25,501-508,558-597 (B)  | 267,556,576,578,582,590   |
| <b>Apc10</b>     | 185    | Doc homology (2-162)<br>X-ray. PDB: 1JHJ                             | IR tail (163-185)<br>De novo   | -  | -   | -   |
| <b>Apc11</b>     | 84     | $\beta$ -strand (1-18)<br>Homology. PDB: 1LDK                        | RING domain (21-84)<br>X-ray. PDB: 4R2Y  | -  | -   | -   |
| <b>Apc12</b>     | 85     | N-term (1-25)<br>X-ray. PDB: 3HYM                                    | -  | -  | 26-85   | -   |
| <b>Apc13</b>     | 74     | De novo  | -  | -  | 40,47,68-79   | -   |
| <b>Apc15</b>     | 121    | De novo  | -  | -  | 57-121  | -   |
| <b>Apc16</b>     | 110    | De novo  | -  | -  | 1-51  | -   |
| <b>Cdh1</b>      | 496    | NTD (1-163)<br>De novo   | WD40 domain (169-470)<br>Homology. PDB: 4BH6   | IR tail (483-496)<br>De novo                                       | 1-41,68-87,109-124,133-145,164-168,471-482  | -   |
| <b>Emi1</b>      | 447    | D box (319-335)<br>X ray/De novo                                     | ZBR (356-372)<br>NMR. PDB: 2M6N  | Linker (356-372), De novo<br>LR tail (431-447), De novo            | 1-318,336-355,373-430   | -   |
| <b>UbcH10</b>    | 179    | UBC domain (30-179)<br>X-ray. PDB: 1I7K                              | -  | -  | 1-29  | -   |
| <b>Ubiquitin</b> | 76     | X-ray. PDB: 4AP4   | -  | -  | -   | -   |

PDB accession numbers are listed. Phosphorylation site data from refs 71–73.

**Extended Data Table 3 | Table comparing APC/C<sup>Cdh1.Hsl1</sup> (ternary APC/C) and dissociation constants for UbcH10 (ref. 4) and UbcH10<sup>LR</sup> (this work).****a. Inter-subunit contact surface area Å<sup>2</sup>**

| Subunit               | Apc1  | Apc2 | Apc3  | Apc4 | Apc5  | Apc6  | Apc7  | Apc8  | Apc10 | Apc11 | Apc12 | Apc13 | Apc15 | Apc16 | Cdh1 | Emi1 |
|-----------------------|-------|------|-------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|------|
| Apc1                  | -     |      |       |      |       |       |       |       |       |       |       |       |       |       |      |      |
| Apc2                  | 2520  | -    |       |      |       |       |       |       |       |       |       |       |       |       |      |      |
| Apc3                  | 360   | 0    | 5746  |      |       |       |       |       |       |       |       |       |       |       |      |      |
| Apc4                  | 0     | 714  | 0     | -    |       |       |       |       |       |       |       |       |       |       |      |      |
| Apc5                  | 8744  | 0    | 0     | 6226 | -     |       |       |       |       |       |       |       |       |       |      |      |
| Apc6                  | 1056  | 0    | 4590  | 0    | 0     | 5244  |       |       |       |       |       |       |       |       |      |      |
| Apc7                  | 0     | 0    | 5354  | 0    | 0     | 0     | 5026  |       |       |       |       |       |       |       |      |      |
| Apc8                  | 4290  | 0    | 0     | 0    | 3722  | 3772  | 0     | 4558  |       |       |       |       |       |       |      |      |
| Apc10                 | 2616  | 0    | 2510  | 0    | 0     | 0     | 0     | 0     | -     |       |       |       |       |       |      |      |
| Apc11                 | 0     | 3402 | 0     | 0    | 0     | 0     | 0     | 0     | 0     | -     |       |       |       |       |      |      |
| Apc11 <sup>RING</sup> | 0     | 666  | 0     | 0    | 0     | 0     | 0     | 0     | 0     | 0     |       |       |       |       |      |      |
| Apc12                 | 246   | 0    | 0     | 0    | 0     | 5970  | 0     | 0     | 0     | 0     | -     |       |       |       |      |      |
| Apc13                 | 0     | 0    | 836   | 0    | 0     | 2068  | 0     | 3240  | 0     | 0     | 0     | -     |       |       |      |      |
| Apc15                 | 0     | 0    | 0     | 224  | 3446  | 340   | 0     | 1082  | 0     | 0     | 0     | 0     | -     |       |      |      |
| Apc16                 | 40    | 0    | 2514  | 0    | 0     | 222   | 1502  | 0     | 0     | 0     | 0     | 0     | 0     | -     |      |      |
| Cdh1                  | 1960  | 0    | 1906  | 0    | 0     | 712   | 0     | 3022  | 0     | 0     | 0     | 0     | 0     | 0     | -    |      |
| Cdh1 <sup>NTD</sup>   | 1960  | 0    | 0     | 0    | 0     | 500   | 0     | 3022  | 0     | 0     | 0     | 0     | 0     | 0     | 0    | 0    |
| Cdh1 <sup>WD40</sup>  | 0     | 0    | 82    | 0    | 0     | 212   | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0    | 0    |
| Cdh1 <sup>IR</sup>    | 0     | 0    | 1824  | 0    | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0    | 0    |
| Emi1                  | 732   | 762  | 0     | 0    | 0     | 0     | 0     | 0     | 704   | 1300  | 0     | 0     | 0     | 0     | 754  | -    |
| Total                 | 24524 | 8064 | 25722 | 7164 | 34010 | 24686 | 11882 | 26708 | 5830  | 4702  | 6216  | 6144  | 5092  | 4278  | 8354 | 4252 |

**b. DALI Results**

| Subunit /domain      | Protein match | RCSB PDB | Z    | RMSD (Å) | Aligned (N) | Res (N) | ID (%) | Subunit /domain | Protein match | RCSB PDB | Z    | RMSD (Å) | Aligned (N) | Res (N) | ID (%) |
|----------------------|---------------|----------|------|----------|-------------|---------|--------|-----------------|---------------|----------|------|----------|-------------|---------|--------|
| Apc1 <sup>WD40</sup> | Sc Cdh1       | 4BH6-E   | 15.9 | 4.0      | 259         | 303     | 8      | Apc6            | Sp Apc6       | 2XPI     | 36.9 | 2.2      | 481         | 520     | 37     |
| Apc1 <sup>PC</sup>   | Sc Rpn2       | 4ADY-B   | 27.8 | 3.2      | 406         | 802     | 15     | Apc7            | Sp Apc6       | 2XPI     | 24.9 | 7.8      | 464         | 518     | 15     |
| Apc2 <sup>NTD</sup>  | Hs Cul3       | 4HXI-B   | 13.8 | 7.4      | 301         | 346     | 7      | Apc8            | Sp Apc6       | 2XPI     | 25.5 | 7.2      | 443         | 518     | 19     |
| Apc2 <sup>CTD</sup>  | Hs Cul1       | 4F52-A   | 23.5 | 2.2      | 237         | 258     | 18     | Apc3A           | Apc3B         | -        |      | 1.6      | 482         | 497     | 100    |
| Apc5 <sup>TPR</sup>  | Sp Apc6       | 2XPI-A   |      |          |             |         |        | Apc6A           | Apc6B         | -        |      | 4.0      | 486         | 503     | 100    |
| Apc3                 | Sp Apc6       | 2XPI-A   | 29.6 | 4.9      | 456         | 518     | 19     | Apc7A           | Apc7B         | -        |      | 2.6      | 483         | 483     | 100    |
|                      |               |          |      |          |             |         |        | Apc8A           | Apc8B         | -        |      | 3.7      | 389         | 490     | 100    |

**c. Surface Plasmon Resonance Data**

| ligand               | Analyte        | $k_{\text{off}}$ (s <sup>-1</sup> ) | $k_{\text{on}}$ (M <sup>-1</sup> s <sup>-1</sup> ) | $K^{\text{eq}}_{\text{d}}$ (nM) | $K^{\text{kin}}_{\text{d}}$ (nM) | Reference  |
|----------------------|----------------|-------------------------------------|--|---------------------------------|----------------------------------|------------|
| UbcH10               | Ternary        | 0.42 ± 0.02                         | 1.78 ± 0.2 × 10 <sup>6</sup>                       | 215                             | 236 ± 29                         | 4          |
| UbcH10 <sup>LR</sup> | Ternary (fast) | 4.0 ± 0.2 × 10 <sup>-2</sup>        | 1.70 ± 0.1 × 10 <sup>6</sup>                       | 31                              | 24 ± 2                           | This study |
|                      | (slow)         | 4.6 ± 0.2 × 10 <sup>-3</sup>        | 1.70 ± 0.1 × 10 <sup>5</sup>                       |                                 | 23 ± 5                           |            |

$K^{\text{kin}}_{\text{d}}$  is the ratio of the observed microscopic constants

a, Inter-subunit contact areas, calculated using CCP4. b, Pairwise structural comparison (DALI) of APC/C subunits. c, Surface plasma resonance data.  $K^{\text{eq}}_{\text{d}}$ , equilibrium dissociation constant;  $K^{\text{kin}}_{\text{d}}$ , kinetic dissociation constant =  $k_{\text{off}}/k_{\text{on}}$ . Standard errors of the fit are listed.

# Galaxies at redshifts 5 to 6 with systematically low dust content and high [C II] emission

P. L. Capak<sup>1,2</sup>, C. Carilli<sup>3,4</sup>, G. Jones<sup>5</sup>, C. M. Casey<sup>6</sup>, D. Riechers<sup>7</sup>, K. Sheth<sup>8</sup>, C. M. Carollo<sup>9</sup>, O. Ilbert<sup>10</sup>, A. Karim<sup>11</sup>, O. LeFevre<sup>10</sup>, S. Lilly<sup>8</sup>, N. Scoville<sup>2</sup>, V. Smolcic<sup>12</sup> & L. Yan<sup>1,2</sup>

**The rest-frame ultraviolet properties of galaxies during the first three billion years of cosmic time (redshift  $z > 4$ ) indicate a rapid evolution in the dust obscuration of such galaxies<sup>1–3</sup>. This evolution implies a change in the average properties of the interstellar medium, but the measurements are systematically uncertain owing to untested assumptions<sup>4,5</sup> and the inability to detect heavily obscured regions of the galaxies. Previous attempts to measure the interstellar medium directly in normal galaxies at these redshifts have failed for a number of reasons<sup>6–9</sup>, with two notable exceptions<sup>10,11</sup>. Here we report measurements of the forbidden C II emission (that is, [C II]) from gas, and the far-infrared emission from dust, in nine typical star-forming galaxies about one billion years after the Big Bang ( $z \approx 5–6$ ). We find that these galaxies have thermal emission that is less than 1/12 that of similar systems about two billion years later, and enhanced [C II] emission relative to the far-infrared continuum, confirming a strong evolution in the properties of the interstellar medium in the early Universe. The gas is distributed over scales of one to eight kiloparsecs, and shows diverse dynamics within the sample. These results are consistent with early galaxies having significantly less dust than typical galaxies seen at  $z < 3$  and being comparable in dust content to local low-metallicity systems<sup>12</sup>.**

We have obtained rest-frame far-infrared (FIR; 3–1,000  $\mu\text{m}$ ) measurements of nine ‘normal’ ( $\sim 1–4$  times average luminosity,  $L_*$ ) galaxies and one low-luminosity quasar at redshifts of  $z = 5–6$ , about one billion years after the Big Bang. These data were taken with an early 20-antenna version of the Atacama Large Millimeter Array (ALMA) in band 7 (1,090–800  $\mu\text{m}$ ) continuum mode with 7.5 GHz of bandwidth,  $\sim 20$  min of integration, and with the compact array yielding a resolution of  $\sim 0.6$  arcsec. One of the four side bands was always centred on the 158  $\mu\text{m}$  [C II] line, which is the dominant FIR cooling line for neutral gas in normal star-forming galaxies<sup>13,14</sup> and hence is a good indicator of galaxy dynamics and the spatial extent of the interstellar medium<sup>14,15</sup> (ISM). The other three spectral bands yield a measurement of the dust continuum emission at  $\lambda \approx 150 \mu\text{m}$ , which is a good indicator of the total infrared luminosity for normal FIR spectral energy distributions (SEDs; see Methods section ‘Derivation of physical parameters’).

The objects we studied were Lyman break galaxies (LBGs) selected from the 2 square degree Cosmic Evolution Survey (COSMOS) field<sup>16</sup>; these LBGs (with one object later discovered to be a low-luminosity quasar) have spectroscopically determined absorption line redshifts measured by the Deep Extragalactic Imaging Multi-Object Spectrograph (DEIMOS) on the W. M. Keck-II Observatory in Hawaii. The objects were selected to be ‘characteristic’ with luminosities between 1 and 4  $L_*$  and with ultraviolet spectral slopes ( $\beta$ ) between  $-1.4$  and  $-0.7$ . Three objects (1, 2 and 10) also meet the selection criteria for

Lyman- $\alpha$  ( $\text{Ly}\alpha$ ) emitters at  $z \approx 5.7$ . When possible, ultraviolet morphologies are measured using the Hubble Space Telescope (HST) Advanced Camera For Surveys (ACS) F814W (0.8  $\mu\text{m}$ ) data<sup>17</sup>. Stellar masses were determined by fitting SED models to existing COSMOS data and new 3–5  $\mu\text{m}$  photometry from the Spitzer-SPLASH survey<sup>18</sup> (see Methods section ‘Optical and NIR data’). Rest-frame infrared luminosities were determined by assuming the range of FIR properties measured at  $z \approx 2–5$  (ref. 19) and have a systematic uncertainty of 0.3 dex due to the uncertainty in the shape of the FIR SED (see Methods section ‘Derivation of physical parameters’).

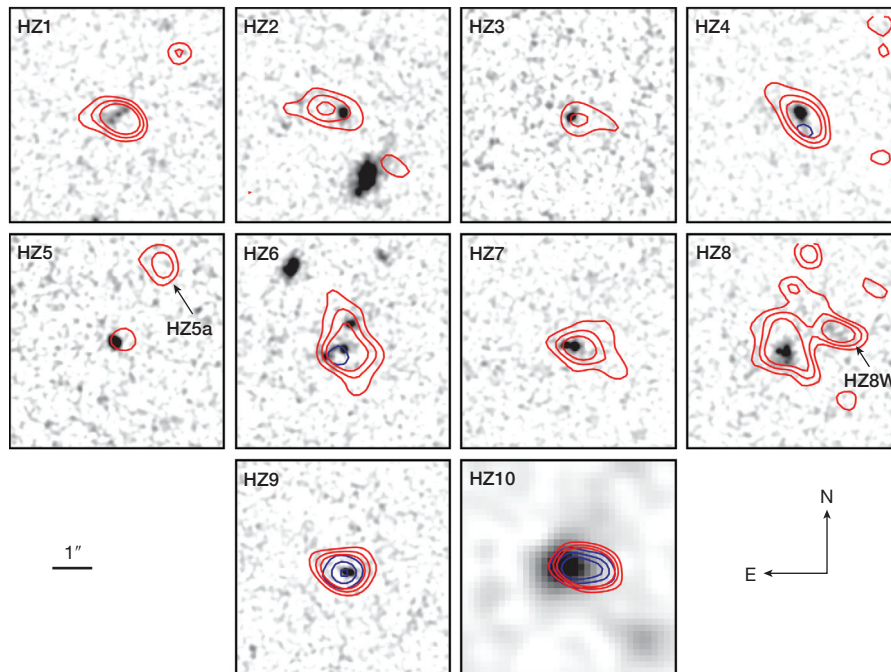
We detect four of nine galaxies in the dust continuum (Fig. 1, Extended Data Table 1), and a mean combination (stacking) of the data on the remaining six objects yields a continuum detection of  $35 \pm 13 \mu\text{Jy}$ . This is less than 1/12 of that expected for systems with similar ultraviolet properties at  $z < 3$  assuming the worst-case systematics in our FIR determination (see Methods section ‘Robustness of upper limits on dust emission’). This can be quantified with the infrared-excess to ultraviolet-slope (IRX- $\beta$ ) relation<sup>4</sup> (Fig. 2), which relates the amount of dust absorption measured in the ultraviolet to the amount of infrared re-emission and is sensitive to both dust properties and column density. We find that the measured IRX and  $\beta$  for our continuum detected objects and A1689-zD1<sup>11</sup> are consistent with the IRX- $\beta$  relation for the Small Magellanic Cloud (SMC)<sup>20</sup>, which has lower-metallicity gas, and less thermal FIR emission<sup>21</sup> than typical galaxies at  $z < 3$ . Our undetected systems appear to have even less IRX than the SMC, but could be consistent given the large systematic uncertainties. Hence, we conclude  $z \approx 6$  galaxies are significantly less dust-obscured than more-evolved systems at  $z < 3$ , confirming previous results<sup>1–3,6</sup>.

The very low IRX values at  $\beta \approx -1.2$  implied by our non-detections are difficult to explain in the context of the very young systems expected at  $z = 5–6$  (ref. 22). The simplest explanation is the systematics in our estimates of  $\beta$  and total infrared luminosity ( $L_{\text{IR}}$ ), which would imply relatively un-obscured systems with  $\sim 45$  K dust (see Methods section ‘Derivation of physical parameters’). But we cannot exclude changes in the dust geometry or rapid disruption of the molecular clouds that could lead to lower IRX values at a given  $\beta$  with more-normal dust temperatures<sup>23</sup>. Higher signal-to-noise near-infrared photometry and shorter wavelength FIR data will ultimately be needed to understand these sources.

A corollary of this result is that ultraviolet derived star-formation rates (SFRs) at high redshift ( $z > 5$ ) should use the SMC-like IRX- $\beta$  relation or assume no dust rather than the currently assumed Meurer IRX- $\beta$  relation. This will decrease the ultraviolet derived SFRs by a typical factor of  $\sim 2–4$  for individual galaxies from those implied by the Meurer relation for similar values of  $\beta$ . The effect on the global

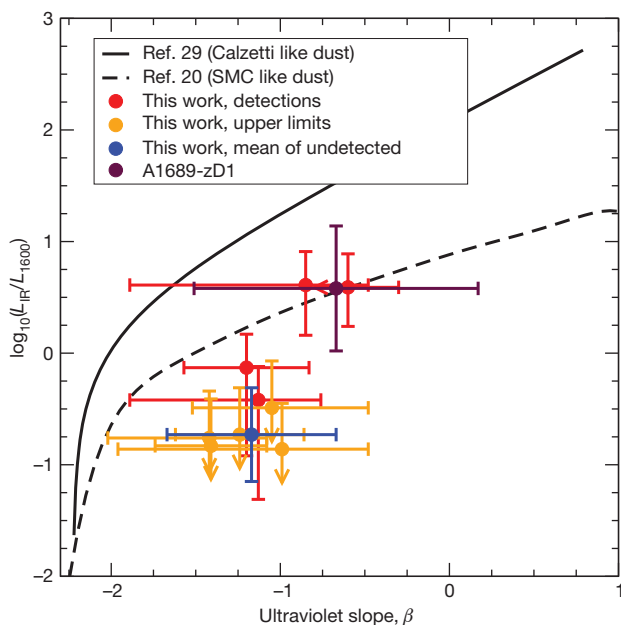
<sup>1</sup>Infrared Processing and Analysis Center (IPAC), 1200 East California Boulevard, Pasadena, California 91125, USA. <sup>2</sup>California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA. <sup>3</sup>National Radio Astronomy Observatory, PO Box 0, Socorro, New Mexico 87801, USA. <sup>4</sup>Astrophysics Group, Cavendish Laboratory, J. J. Thomson Avenue, Cambridge CB3 0HE, UK. <sup>5</sup>New Mexico Institute of Mining and Technology, 801 Leroy Place, Socorro, New Mexico 87801, USA. <sup>6</sup>Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, Texas 78712, USA. <sup>7</sup>Department of Astronomy, Cornell University, 220 Space Sciences Building, Ithaca, New York 14853, USA. <sup>8</sup>National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, Virginia 22903, USA. <sup>9</sup>Institute for Astronomy, ETH Zurich, CH-8093 Zurich, Switzerland. <sup>10</sup>Aix Marseille Université, CNRS, LAM (Laboratoire d’Astrophysique de Marseille), UMR 7326, 13388 Marseille, France. <sup>11</sup>Argelander-Institut für Astronomie, Auf dem Hügel 71, D-53121 Bonn, Germany. <sup>12</sup>Physics Department, University of Zagreb, Bijenička cesta 32, 10002 Zagreb, Croatia.





**Figure 1 | Optical, [C II] and continuum maps of the sources HZ1–HZ10.** The [C II] line detections (red contours) and weak  $\sim 158 \mu\text{m}$  FIR continuum detections (blue contours) are shown with the rest-frame ultraviolet images as the background. The images are  $5'' \times 5''$  (scale bar at bottom left) and the contours are 2, 6 and  $10 \sigma$  with [C II] line profiles for each source shown in

Fig. 3. The background images are from HST-ACS in the F814W<sup>17</sup> band where the morphologies will be affected by Ly $\alpha$ , except for HZ10, which is Subaru z' band. All objects are detected in [C II], showing that a large amount of gas is present in these systems, but only four are detected in continuum.

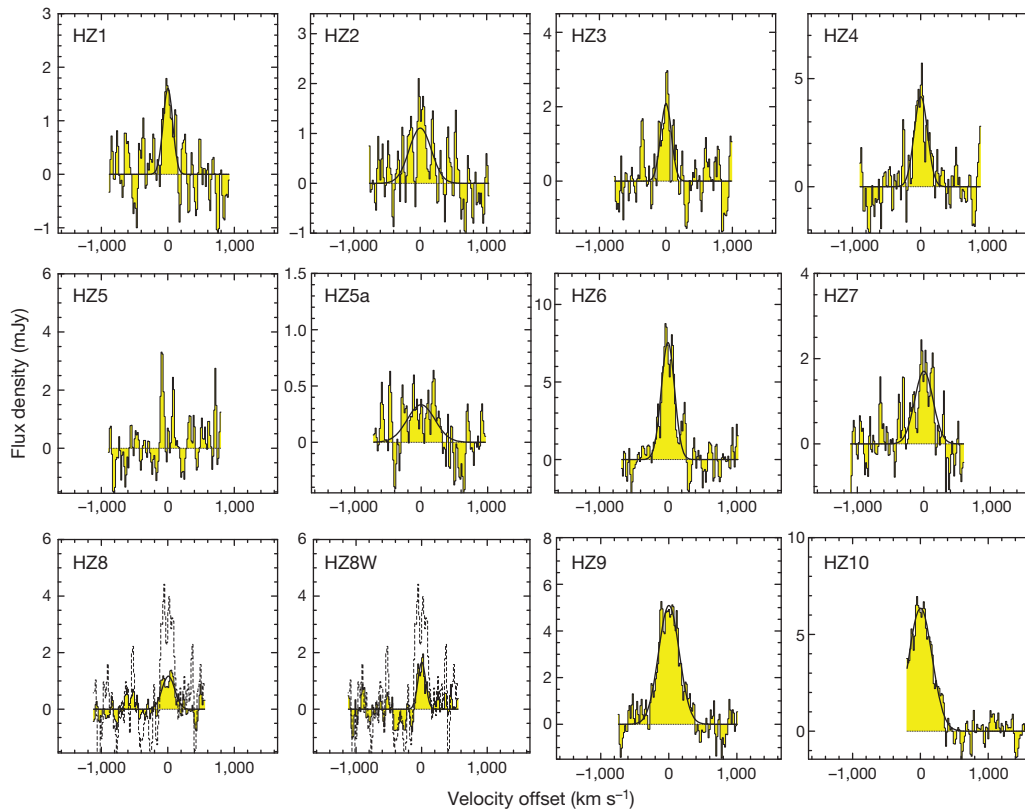


**Figure 2 | IRX– $\beta$  measurements of  $z > 5$  objects.** The deficit of infrared emission in our sample is evident in the presented infrared excess versus ultraviolet slope (IRX– $\beta$ ) relation when compared with models. Detections are indicated in red, upper limits in orange, the mean IRX ratio (obtained by combining undetected sources) in blue, and A1689-zD1<sup>11</sup> in maroon. Error bars are  $1\sigma$ , and include standard measurement error and systematic uncertainty added in quadrature. The Meurer<sup>29</sup> relation, which assumes Calzetti-like dust and is consistent with typical galaxies at  $z < 3$ , is shown as a black solid line, while a model for lower-metallicity SMC-like dust model<sup>20</sup> is shown as a dashed line.

star-formation history is much smaller,  $<40\%$ , because the majority of star-formation is in low luminosity ( $<L_*$ ) galaxies that were already assumed to have little or no dust extinction (see Methods section ‘Effects of evolving dust on the global star-formation history’, Extended Data Fig. 4).

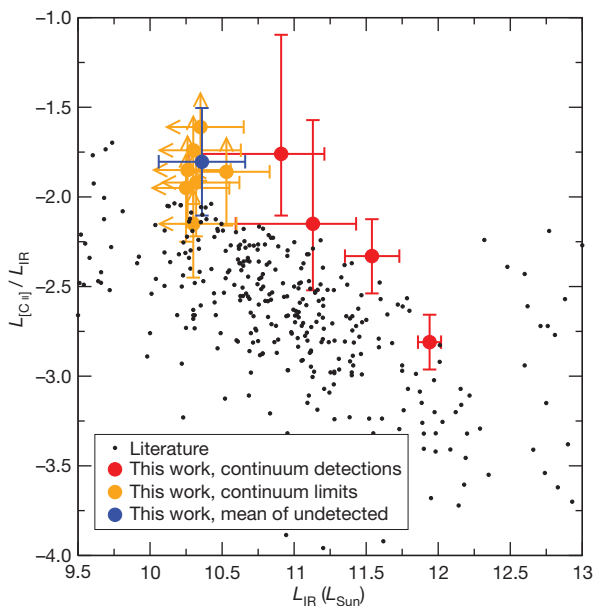
In contrast to the dust emission, we find  $>3\sigma$  detections of the [C II] line in all nine normal galaxies (Fig. 3). The line emission is spectrally resolved in all cases, with [C II] velocity dispersions of  $63\text{--}163 \text{ km s}^{-1}$ , and marginally resolved at our spatial resolution of  $\sim 0.5\text{--}0.9$  arcsec, indicating galaxies with dynamic masses ( $M_{\text{dyn}}$ ) of  $\sim 10^9\text{--}10^{11}$  solar masses ( $M_\odot$ ; see Methods sections ‘Reduction of ALMA data’ and ‘Derivation of physical parameters’). We also detect two optically faint [C II] emitters at redshifts consistent with the targeted objects. HZ5a is detected near HZ5 at a redshift consistent with the in-falling gas seen in the optical spectra of HZ5<sup>24</sup>. HZ8W corresponds to an optically faint companion to HZ8 and has a similar redshift. Taken together, the direct and serendipitous detections suggest ubiquitous and enhanced [C II] emission in early galaxies similar to that seen in local low-metallicity systems<sup>12</sup> (Fig. 4).

The [C II] enhancement in local systems is caused by a lower dust-to-gas ratio which allows the ultraviolet radiation field to penetrate a larger volume of molecular cloud<sup>12</sup>. Our significantly lower IRX values and enhanced [C II]/FIR ratios would suggest a similar effect is happening in high redshift galaxies. But other possible causes of the transition in obscuration properties with redshift have been suggested<sup>25</sup>: evolution in metal abundances that changes the intrinsic ultraviolet slope; changes in the dust properties; and differences in the dust geometry. The systems in this study were selected to have broad ultraviolet absorption features in their spectra that indicate a relatively homogeneous metal abundance of  $\sim 0.25$  times the solar value. At this metallicity, the ultraviolet spectral slopes are expected to be similar to those of solar metallicity



**Figure 3 | [C II] line profiles for the sources.** The line velocity profiles of our [C II] measurements are shown for our 10 sources and two blind detections, HZ5a and HZ8W. All sources but HZ5, a quasar, are clearly detected. For

objects HZ8 and HZ8W, a dotted line showing the integrated flux over both sources is also plotted. The consistency between the integrated and individual profiles indicates they are at similar redshifts.



**Figure 4 |  $L_{[\text{C II}]}$  over  $L_{\text{IR}}$  ratio as a function of  $L_{\text{IR}}$ .** Our sample has a clear excess of [C II] emission relative to IR emission when compared with normal lower-redshift systems ( $z = 0-3$ )<sup>6,12,15,30</sup>. Detections are indicated in red, upper limits in orange, and the mean of undetected sources is shown in blue. The high  $L_{[\text{C II}]} / L_{\text{IR}}$  ratio we measure indicates a lower dust-to-gas ratio, more-diffuse clouds, and a more-diffuse ultraviolet radiation field than seen in normal  $z \approx 0$  galaxies<sup>12,13,15,30</sup>. Error bars are  $1\sigma$ , including standard measurement error and systematics added in quadrature.

systems<sup>26</sup>. Furthermore, the population from which these objects were selected has dust attenuation properties similar to those of lower-redshift objects<sup>27</sup>. So changes in the dust properties that would be measurable in the attenuation curve are probably not the primary factor. When compared with objects of the same stellar masses, our sample has physical properties (gas velocity dispersions, sizes, and SFRs) that imply geometries consistent with  $z \approx 1-3$  galaxies that exhibit dust properties like those found in the Meurer model<sup>28</sup>. Changes in the average cloud geometry are therefore not obvious, but we cannot fully exclude this possibility as it would explain the observed IRX properties. So we conclude that a significant decrease in the dust-to-gas ratio is the most likely explanation for the evolution in extinction properties.

These results are seemingly at odds with previous measurements<sup>6-9</sup> that failed to detect [C II] emission at high redshift. Our sample contains objects with both emission and absorption of Ly $\alpha$ , and the [C II] emission strength is un-correlated with the Ly $\alpha$  properties (Extended Data Table 1), which means that the use of Ly $\alpha$  redshifts is not the cause. The most likely explanation for the lack of detections is observational uncertainty in the expected line flux, line width, and frequency. Even obscuration uncorrected ultraviolet SFRs will overestimate the expected FIR continuum flux, because less than 20% of the star formation is leading to FIR emission in our sample<sup>6</sup> (Fig. 2). This means the FIR emission of galaxies will be more difficult to detect than previously expected based on  $z < 3$  scaling relations (see Methods sections ‘Robustness of upper limits on dust emission’ and ‘Quantifying [C II] non-detections in the literature’), so longer integration times are required. That said, these data show [C II] is enhanced and readily detectable with ALMA at  $z \approx 6$  if moderate integration times are used. Finally, we note that the [C II] lines are resolved in velocity space, so single channel sensitivity should not be used for

limits (see Methods section ‘Quantifying [C II] non-detections in the literature’).

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 20 November 2014; accepted 23 April 2015.**

- Madau, P. & Dickinson, M. Cosmic star-formation history. *Annu. Rev. Astron. Astrophys.* **52**, 415–486 (2014).
- Bouwens, R. J. *et al.* UV-continuum slopes at  $z \sim 4$ –7 from the HUDF09+ERS+CANDELS observations: discovery of a well-defined UV color-magnitude relationship for  $z \geq 4$  star-forming galaxies. *Astrophys. J.* **754**, 83 (2012).
- Finkelstein, S. L. *et al.* Candelas: the evolution of galaxy rest-frame ultraviolet colors from  $z = 8$  to 4. *Astrophys. J.* **756**, 164 (2012).
- Reddy, N. A., Erb, D. K., Pettini, M., Steidel, C. C. & Shapley, A. E. Dust obscuration and metallicity at high redshift: new inferences from UV, H $\alpha$ , and 8  $\mu$ m observations of  $z \sim 2$  star-forming galaxies. *Astrophys. J.* **712**, 1070–1091 (2010).
- Boquien, M. *et al.* The IRX- $\beta$  relation on subgalactic scales in star-forming galaxies of the Herschel Reference Survey. *Astron. Astrophys.* **539**, A145 (2012).
- Ota, K. *et al.* ALMA observation of 158  $\mu$ m [C II] line and dust continuum of a  $z = 7$  normally star-forming galaxy in the epoch of reionization. *Astrophys. J.* **792**, 34 (2014).
- Ouchi, M. *et al.* An intensely star-forming galaxy at  $z \sim 7$  with low dust and metal content revealed by deep ALMA and HST observations. *Astrophys. J.* **778**, 102 (2013).
- Maiolino, R. *et al.* The assembly of “normal” galaxies at  $z=7$  probed by ALMA. Preprint at <http://arXiv.org/abs/1502.06634> (2015).
- Schaefer, D. *et al.* New constraints on dust emission and UV attenuation of  $z = 6.5$ –7.5 galaxies from millimeter observations. *Astron. Astrophys.* **574**, A19 (2015).
- Riechers, D. A. *et al.* ALMA imaging of gas and dust in a galaxy protocluster at redshift 5.3: [C II] emission in “typical” galaxies and dusty starbursts  $\sim 1$  billion years after the Big Bang. *Astrophys. J.* **796**, 84 (2014).
- Watson, D. *et al.* A dusty, normal galaxy in the epoch of reionization. *Nature* **519**, 327–330 (2015).
- Israel, F. P. & Maloney, P. R. C<sup>+</sup> emission from the Magellanic Clouds. II. [C II] maps of star-forming regions LMC-N 11, SMC-N 66, and several others. *Astron. Astrophys.* **531**, A19 (2011).
- Malhotra, S. *et al.* Infrared Space Observatory measurements of [C II] line variations in galaxies. *Astrophys. J.* **491**, L27–L30 (1997).
- Carilli, C. L. & Walter, F. Cool gas in high-redshift galaxies. *Annu. Rev. Astron. Astrophys.* **51**, 105–161 (2013).
- Díaz-Santos, T. *et al.* Extended [C II] emission in local luminous infrared galaxies. *Astrophys. J.* **788**, L17 (2014).
- Leauthaud, A. *et al.* Weak gravitational lensing with COSMOS: galaxy selection and shape measurements. *Astrophys. J.* **172** (Suppl.), 219–238 (2007).
- Koekemoer, A. M. *et al.* The COSMOS survey: Hubble Space Telescope Advanced Camera for Surveys observations and data processing. *Astrophys. J.* **172** (Suppl.), 196–202 (2007).
- Speagle, J. S., Steinhardt, C. L., Capak, P. L. & Silverman, J. D. A highly consistent framework for the evolution of the star-forming “main sequence” from  $z \sim 0$ –6. *Astrophys. J.*, Suppl. 214, 15–67 (2014).
- Casey, C. M. Far-infrared spectral energy distribution fitting for galaxies near and far. *Mon. Not. R. Astron. Soc.* **425**, 3094–3103 (2012).
- Pettini, M. *et al.* Infrared observations of nebular emission lines from galaxies at  $z \sim 3$ . *Astrophys. J.* **508**, 539–550 (1998).
- Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U. & Wolff, M. J. A quantitative comparison of the Small Magellanic Cloud, Large Magellanic Cloud, and Milky Way ultraviolet to near-infrared extinction curves. *Astrophys. J.* **594**, 279–293 (2003).
- Kong, X., Charlot, S., Brinchmann, J. & Fall, S. M. Star formation history and dust content of galaxies drawn from ultraviolet surveys. *Mon. Not. R. Astron. Soc.* **349**, 769–778 (2004).
- Charlot, S. & Fall, S. M. A simple model for the absorption of starlight by dust in galaxies. *Astrophys. J.* **539**, 718–731 (2000).
- Capak, P. L. *et al.* A massive protocluster of galaxies at a redshift of  $z \sim 5.3$ . *Nature* **470**, 233–235 (2011).
- Reddy, N. *et al.* GOODS-Herschel measurements of the dust attenuation of typical star-forming galaxies at high redshift: observations of ultraviolet-selected galaxies at  $z \sim 2$ . *Astrophys. J.* **744**, 154 (2012).
- Leitherer, C., Tremonti, C. A., Heckman, T. M. & Calzetti, D. An ultraviolet spectroscopic atlas of local starbursts and star-forming galaxies: the legacy of FOS and GHRS. *Astron. J.* **141**, 37 (2011).
- Scoville, N. *et al.* Dust attenuation in high redshift galaxies — ‘diamonds in the sky’. *Astrophys. J.* **800**, 108 (2015).
- Förster Schreiber, N. M. *et al.* The SINS survey: SINFONI integral field spectroscopy of  $z \sim 2$  star-forming galaxies. *Astrophys. J.* **706**, 1364–1428 (2009).
- Meurer, G. R., Heckman, T. M. & Calzetti, D. Dust absorption and the ultraviolet luminosity density at  $z \sim 3$  as calibrated by local starburst galaxies. *Astrophys. J.* **521**, 64–80 (1999).
- Díaz-Santos, T. *et al.* Explaining the [C II] 157.7  $\mu$ m deficit in luminous infrared galaxies — first results from a Herschel/PACS study of the GOALS sample. *Astrophys. J.* **774**, 68 (2013).

**Acknowledgements** Support for this work was provided by NASA through an award issued by JPL/Caltech. We thank the ALMA staff for facilitating the observations and aiding in the calibration and reduction process. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. This work is based in part on observations made with the Spitzer Space Telescope and the W.M. Keck Observatory, along with archival data from the NASA/ESA Hubble Space Telescope, the Subaru Telescope, the Canada-France-Hawaii-Telescope and the ESO Vista telescope obtained from the NASA/IPAC Infrared Science Archive. V.S. acknowledges funding by the European Union’s Seventh Framework programme under grant agreement 337595 (ERC Starting Grant, ‘CoSMass’).

**Author Contributions** P.L.C. proposed and carried out the observations, conducted the analysis in this paper, and authored the majority of the text. C.C., G.J. and K.S. carried out the reduction and direct analysis of the ALMA data. C.M.C. consulted on the spectral energy distribution fitting and interpretation of the data, and also conducted a blind test of the FIR luminosity, [C II] line luminosity, and  $\beta$  measurements, along with testing for sample selection effects. D.R. conducted the spectral line analysis and carried out an independent blind check of the ALMA data reduction. O.I. carried out the spectral energy distribution fitting and consulted on their interpretation. C.M.C., A.K., O.L., S.L., N.S., V.S. and L.Y. contributed to the overall interpretation of the results and various aspects of the analysis.

**Author Information** This paper makes use of ALMA data: ADS/JAO.ALMA#2012.1.00523.S. ALMA Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.L.C. ([capak@astro.caltech.edu](mailto:capak@astro.caltech.edu)).



## METHODS

**Cosmology.** Throughout this paper we assume a simplified cosmology of  $\Omega_M = 0.3$ ,  $\Omega_V = 0.7$  and  $H_0 = 0.7$  when calculating physical parameters. These values approximate the 2013 Planck<sup>31</sup> and 9-year WMAP<sup>32</sup> values, and are commonly assumed in the current literature to make comparisons easier over time as the exact measured values evolve.

**Reduction of ALMA data.** The ALMA observatory staff, as part of standard data processing and delivery, performed initial data calibration. The calibrated visibility data were then re-analysed, and additional flagging of bad time periods in the data and bad channels was performed. The data were re-imaged using the CASA Briggs flexible weighting of the u-v visibility data with Robust = 1 to create continuum, moment zero, and channel maps. Gaussian fitting of the spectral and spatial data was performed using the CASA viewer and CASA fit tool.

The centroid frequency of these models was compared to the rest frequency of [C II] to measure the redshift, which was compared to the previously determined redshifts from optical spectroscopy. Since these measures were identical, within our precision, the reference frequency of each cube was set to the centroid frequency, allowing the FWHM of the Gaussian to be read in  $\text{km s}^{-1}$  and the line flux in  $\text{Jy km s}^{-1}$ . The [C II] line contribution was subtracted from the measured continuum fluxes.

The measurements derived from the ALMA data are given in Extended Data Table 1 and Gaussian fits done with CASA to the image components are given in Extended Data Tables 2 and 4 for the [C II] and continuum data, respectively. The fits were not constrained by the beam size, and so they can be slightly smaller due to the data errors. The errors quoted are those produced by the CASA Gaussian fitting routine and are sensitive to the box size chosen for the fit, especially in cases of low signal to noise. Deconvolved size measurements for the [C II] lines are given in Extended Data Table 3. Only objects 3 and 5a are unresolved in [C II]. In contrast, only object 10 is resolved in the continuum measurement. The baseline configuration used will resolve out spatial scales larger than  $5\text{--}7''$  ( $29\text{--}43$  kpc) but none of our sources appear to have emission on that large a scale.

**Optical and NIR data.** These were obtained from the COSMOS photometric redshift catalogue<sup>33</sup> augmented with additional data from the Spitzer-Large Area Survey with Hyper-Suprime-Cam (SPLASH)<sup>34</sup>. Sizes were taken from the COSMOS Hubble Space Telescope Advanced Camera for Surveys (HST-ACS) weak lensing catalogue<sup>16</sup> and are given in Extended Data Table 5. Object 10 falls outside the area covered by HST-ACS.

**Derivation of physical parameters.** Since only one FIR data point was available, total infrared ( $3\text{--}1,100\ \mu\text{m}$ ) luminosity was derived using the range of grey body models covering the full range of objects observed at  $z \approx 2\text{--}5$ <sup>19</sup>. The grey body model parameters,  $\alpha$  the blue power-law,  $\beta$  the long wavelength slope, and  $T$  the black body temperature, ranged between  $\alpha = 1.5\text{--}2.5$ ,  $\beta = 1.2\text{--}2.0$  and  $T = 25\text{--}45$  K. The range of allowed infrared luminosities was determined by scaling these grey bodies to the observed rest frame  $158\ \mu\text{m}$  flux. This range of parameters resulted in a systematic uncertainty of 0.3 dex in the derived luminosity. This relatively small error is due to the fact that  $158\ \mu\text{m}$  is on the flat part of the SED close to the emission peak for the range of observed physical parameters. The derived infrared luminosities are given in Extended Data Table 5.

The rest frame ultraviolet properties were determined by fitting a power-law to the rest frame  $0.13\text{--}0.3\ \mu\text{m}$  photometry from the COSMOS and Ultra-Vista surveys. This power-law fit was then used to estimate the ultraviolet spectral slope ( $\beta$ ) and the luminosity at  $0.16\ \mu\text{m}$ , where  $L_{UV} = \nu_{1,600} L_{\nu_{1,600}}$ . The power-law fits appear to be biased high by the bluest bands in several cases which may be due to the intrinsic SED flattening at short wavelengths<sup>35,36</sup>. To account for this possible systematic, we also fit the slope using only the rest-frame  $0.15\text{--}0.3\ \mu\text{m}$  photometry and include the shifts in our measurement error bars. Finally, simulations<sup>2</sup> of observational biases from blending and signal-to-noise indicate that an additional  $\pm 0.3$  systematic error should be included. These estimates are tabulated in Extended Data Table 5.

For consistency, we re-analysed A1689-zD1<sup>11</sup> with the same methods used for our sample. We find  $L_{IR} = (6.8 \pm 1.2) \times 10^{10} L_{\odot}$ ,  $L_{UV} = (1.8 \pm 0.4) \times 10^{10} L_{\odot}$  and  $\beta = -0.67 \pm 0.78$  assuming a lensing magnification of 9.3 versus  $L_{IR} = (6.2 \pm 0.8) \times 10^{10} L_{\odot}$  and  $L_{UV} = 1.8 \times 10^{10} L_{\odot}$ . No value was published for  $\beta$  derived from the photometry, but our analysis would not have produced a consistent  $L_{UV}$  if the two analyses of  $\beta$  were inconsistent because  $L_{UV}$  is derived from the  $\beta$  fit.

Stellar masses were derived using data from the Spitzer Large Area Survey with Hyper-Suprime-Cam (SPLASH) in the same way described in previous work<sup>33</sup>. In brief, the COSMOS and SPLASH  $0.1\text{--}5\ \mu\text{m}$  photometry was fitted with Bruzual and Charlot templates<sup>37</sup> with a Chabrier IMF allowing for a range of emission line strengths and extinction laws. The estimated stellar masses are given in Extended Data Table 5.

Dynamical masses were estimated based on the [C II] velocity dispersion ( $\sigma_{[C II]}$ ) and size estimates in Extended Data Table 3. We used the method outlined in Wang *et al.*<sup>38</sup> to estimate the dynamical masses. In this approximation  $M_{\text{dyn}} = 1.16 \times 10^5 V_{\text{cir}}^2 D$ , where  $V_{\text{cir}}$  is the circular velocity in  $\text{km s}^{-1}$  and  $D$  is the diameter in kpc.  $V_{\text{cir}} = 1.763 \sigma_{[C II]} / \sin(i)$ , and  $i$ , the disk inclination angle, is approximated by the axis ratio of the object as  $i = \cos^{-1}(b/a)$ .

This method yields an approximate answer but is very uncertain because of our low spatial resolution that leads to a large uncertainty in both  $i$  and the object size. The errors quoted in Extended Data Table 5 include the error in the semi-major axis and the velocity dispersion and assumes  $\sin(i) = 0.45\text{--}1$ . This range of  $\sin(i)$  approximates the range of values from a dispersion dominated system to an edge-on disk, but could be significantly larger.

We find the dynamical masses are typically a factor of  $\sim 3$  greater than the stellar masses (see Extended Data Fig. 3). This is higher than the factor of  $1.2\text{--}1.7$  found at  $z \approx 1\text{--}3$  from similar H $\alpha$  based dynamical mass estimates<sup>28</sup>, but consistent given the very large errors in our measurements.

**Robustness of upper limits on dust emission.** The conclusion that the dust properties, and specifically the dust mass and dust to gas ratio, of galaxies are evolving relies on two quantities: the measured deficit of rest-frame  $158\ \mu\text{m}$  flux, and the models that translate this into a deficit of thermal emission. Here we quantify the robustness of this result.

If we assume a Meurer *et al.* like IRX- $\beta$  relation and the mean of our grey body models, the predicted mean ALMA flux of the undetected sources is  $833\text{--}7,610\ \mu\text{Jy}$  (given the uncertainty in  $\beta$ ), which is  $>61\sigma$  discrepant from the measured mean of  $35 \pm 13\ \mu\text{Jy}$ . Individual sources are discrepant from the Meurer *et al.* relation by  $>17\text{--}490\sigma$  except for objects 9 and 10 which are consistent given the large errors in  $\beta$ . If we assume the maximal range of the systematic error,  $45$  K dust,  $\beta = 1.5$ ,  $\alpha = 1.2$ , the predicted mean ALMA flux is  $447\text{--}4,087\ \mu\text{Jy}$ , which is  $>32\sigma$  discrepant from the measured mean. Finally we consider an extreme model with  $100$  K dust,  $\beta = 1.5$ ,  $\alpha = 1.2$ ; we predict an ALMA flux of  $78\text{--}710\ \mu\text{Jy}$ , which is  $3\sigma$  discrepant from the measured mean. In this extreme case, objects 4, 9 and 10 are consistent with the Meurer *et al.* relation, and the remaining sources are discrepant at the  $>1\text{--}30\sigma$  level.

We also performed a completeness analysis following Casey *et al.*<sup>39</sup> and find the low dust emission is unlikely to be due to selection effects in our ultraviolet sample. However, we note the result only holds for ultraviolet selected samples. The general galaxy population may contain heavily obscured objects that may not be easily found by ultraviolet selected surveys.

**Quantifying [C II] non-detections in the literature.** Our sample predicts a peak [C II] line flux of  $\sim 0.2\text{--}0.7\ \text{mJy}$  for an  $L^*$  (ref. 40) galaxy with an SFR  $\approx 6.5$  at  $z \approx 6\text{--}7$ , below the detection limit of most previous studies of  $z \approx 6\text{--}7$  objects<sup>6–9</sup>.

In addition to sensitivity, we note Ly $\alpha$  redshifts are systematically offset by  $\sim 400\ \text{km s}^{-1}$  and up to  $1,200\ \text{km s}^{-1}$  from the [C II] (and systemic absorption line) redshifts<sup>41</sup>, which is a significant fraction of the ALMA band pass.

Finally, caution must be used when interpreting the limits given in the SFR- $L_{[C II]}$  relation because the estimated line luminosity limits scales as the square root of the assumed line width. A  $36\text{--}100\ \text{km s}^{-1}$  FWHM channel width<sup>6–9</sup> is typically used for these limits, but our measured line widths in Extended Data Table 1 are significantly broader. If we use our measured line widths this would predict for [C II] a line limit  $0.2\text{--}1.0$  dex higher than those typically quoted, making them consistent with the SFR- $L_{[C II]}$  relation.

**Comparison of gas and SFR properties to general redshift scaling relations.** This sample is enhanced in  $L_{[C II]}$  relative to  $L_{IR}$  compared with lower-redshift samples<sup>6</sup> (Fig. 4). These sources have  $\sim 1\%$  of their luminosity emitted in the [C II], consistent with low metallicity and extended star forming objects<sup>12,15</sup>.

This sample appears to follow the local SFR- $L_{[C II]}$  relation if the SFR is determined by summing the ultraviolet and FIR measurements<sup>6</sup> rather than using a dust corrected ultraviolet estimate (Extended Data Fig. 1). However, if an ultraviolet derived dust correction is used, the points shift  $\sim 0.5\text{--}1$  dex higher.

Finally we turn to the stellar-mass-SFR or ‘main-sequence’ relation (Extended Data Fig. 2)<sup>18</sup>. We find our objects generally agree with this relation, but several objects fall below the relation. Specifically, objects HZ1, HZ2 and HZ3 are below it, but HZ1 has an unusually low dynamical to stellar mass ratio, indicating the stellar mass may be over-estimated.

**Effects of evolving dust on the global star-formation history.** Our results indicate the properties of dust and the amount of extinction evolve significantly between  $z \approx 3$ , where it is well measured, and  $z \approx 5\text{--}6$ , where we measure it. To estimate the effect on the global star-formation history, we adopt the luminosity functions and ultraviolet slope measurements from the series of papers by Bouwens *et al.*<sup>2,42</sup>, which are largely supported by other work<sup>43</sup>. We then assume the dust correction can be described by the Meurer<sup>29</sup> curve at  $z < 4$ , the SMC<sup>20</sup> curve at  $z \approx 5$ , and the SMC curve scaled to our measured IRX at  $z \approx 6\text{--}8$ . The results of this analysis are shown in Extended Data Fig. 4 and result in

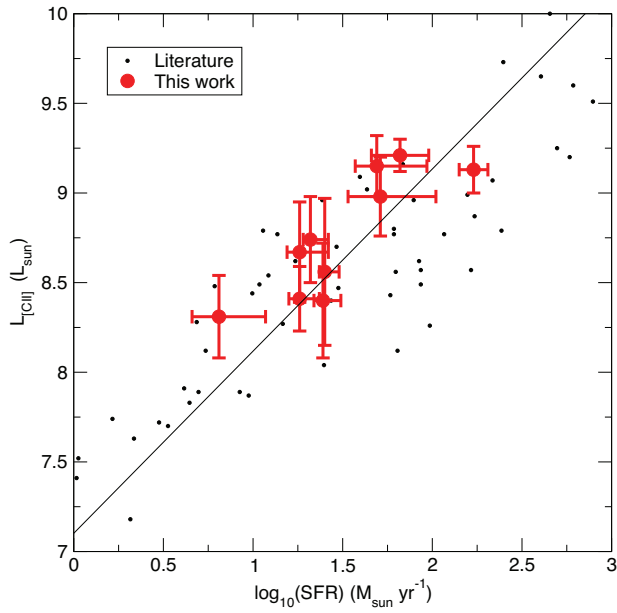
~30–40% less star-formation at  $z \approx 6$  than previously thought. For consistency with the literature, the SFR density is shown assuming a Salpeter IMF; a factor of 0.68 needs to be applied to convert to a Chabrier IMF.

**Sample size.** No statistical methods were used to predetermine sample size.

**Randomization.** The experiments were not randomized.

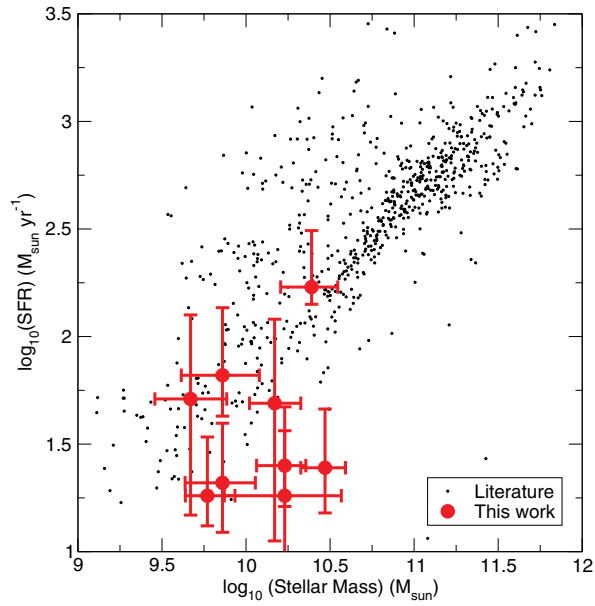
**Blinding.** The investigators were not blinded to allocation during experiments and outcome assessment.

31. Planck Collaboration. Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.* **571**, A16 (2014).
32. Hinshaw, G. *et al.* Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: cosmological parameter results. *Astrophys. J.* **208** (Suppl.), 19 (2013).
33. Ilbert, O. *et al.* Mass assembly in quiescent and star-forming galaxies since  $z \approx 4$  from UltraVISTA. *Astron. Astrophys.* **556**, A55 (2013).
34. Steinhardt, C. L. *et al.* Star formation at  $4 < z < 6$  from the Spitzer Large Area Survey with Hyper-Suprime-Cam (SPLASH). *Astrophys. J.* **791**, L25 (2014).
35. Leitherer, C., Li, I.-H., Calzetti, D. & Heckman, T. M. Global far-ultraviolet (912–1800 Å) properties of star-forming galaxies. *Astrophys. J.* **140** (Suppl.), 303–329 (2002).
36. Buat, V., Burgarella, D., Deharveng, J. M. & Kunth, D. Spectral energy distributions of starburst galaxies in the 900–1200 Å range. *Astron. Astrophys.* **393**, 33–42 (2002).
37. Bruzual, G. & Charlot, S. Stellar population synthesis at the resolution of 2003. *Mon. Not. R. Astron. Soc.* **344**, 1000–1028 (2003).
38. Wang, R. *et al.* Star formation and gas kinematics of quasar host galaxies at  $z \sim 6$ : new insights from ALMA. *Astrophys. J.* **773**, 44 (2013).
39. Casey, C. M. *et al.* Are dusty galaxies blue? Insights on UV attenuation from dust-selected galaxies. *Astrophys. J.* **796**, 95 (2014).
40. Bouwens, R. J. *et al.* UV luminosity functions at redshifts  $z \sim 4$  to  $z \sim 10$ : 10,000 galaxies from HST legacy fields. *Astrophys. J.* **803**, 34 (2015).
41. Shapley, A. E., Steidel, C. C., Pettini, M. & Adelberger, K. L. Rest-frame ultraviolet spectra of  $z \sim 3$  Lyman break galaxies. *Astrophys. J.* **588**, 65–89 (2003).
42. Bouwens, R. J. *et al.* Very blue UV-continuum slope  $\beta$  of low luminosity  $z \sim 7$  galaxies from WFC3/IR: evidence for extremely low metallicities? *Astrophys. J.* **708**, L69–L73 (2010).
43. Rogers, A. B. *et al.* The colour distribution of galaxies at redshift five. *Mon. Not. R. Astron. Soc.* **440**, 3714–3725 (2014).
44. De Looze, I., Baes, M., Bendo, G. J., Cortese, L. & Fritz, J. The reliability of [C ii] as an indicator of the star formation rate. *Mon. Not. R. Astron. Soc.* **416**, 2712–2724 (2011).



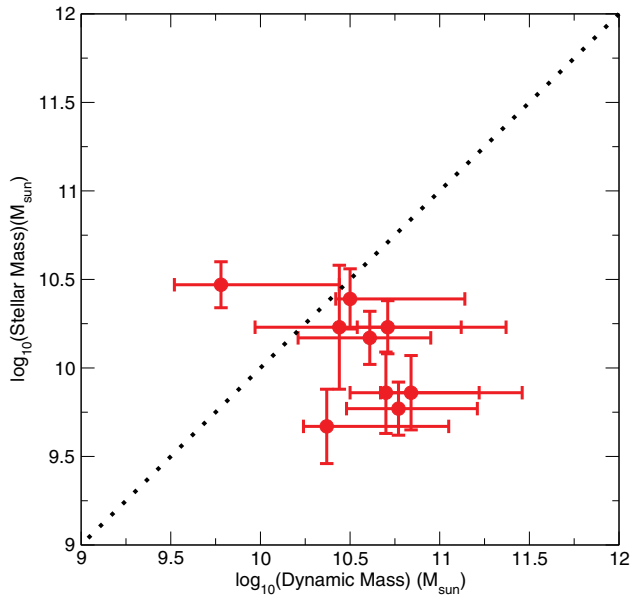
**Extended Data Figure 1 | [C II] flux as a function of total star-formation rate, SFR.** Our sample is consistent with the low- $z$  SFR– $L_{[\text{C II}]}$  relation from the literature at  $z = 0$ –6 (ref. 6). The best fit to the literature points from ref. 44 is indicated with a solid line. The SFRs are derived using the UV + FIR method and a Chabrier IMF<sup>18</sup>. The dust-corrected ultraviolet estimates with a Meurer relation typically used at these redshifts would increase the estimated SFR by a factor of  $\sim 2$ –10 (0.5–1 dex), leading to over-estimates of the expected [C II] flux. The  $L_{[\text{C II}]}$  error bars are  $1\sigma$  standard measurement error, while the SFR errors are  $1\sigma$  from a combination of measurement error in  $L_{\text{IR}}$  and  $L_{\text{UV}}$  converted to star formation added in quadrature.



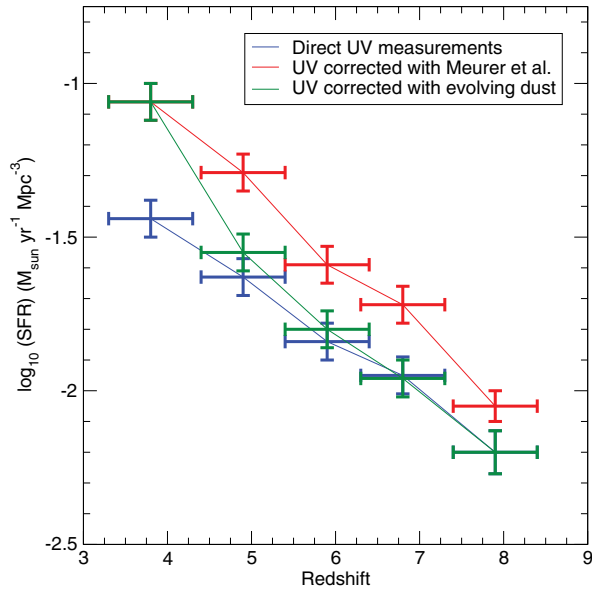


**Extended Data Figure 2 | Star-formation rate as a function of stellar mass.**

Our objects are generally consistent with the stellar mass-SFR or ‘main sequence’ of star forming galaxies at  $z \approx 5-6$  (ref. 18). Three objects fall below the relation, namely HZ1, HZ2 and HZ3, but the stellar mass of HZ1 may be over-estimated (‘Derivation of physical parameters’ in Methods). Mass errors reflect the  $1\sigma$  range of masses allowed by the SED model fits including emission line strength as a free parameter. SFR errors are  $1\sigma$  from a combination of measurement error in  $L_{\text{IR}}$ , systematics in  $L_{\text{IR}}$ , and measurement error in  $L_{\text{UV}}$  converted to star formation added in quadrature.



**Extended Data Figure 3 | Dynamical mass compared with stellar mass.** A comparison of the dynamical masses estimated from the [C II] line and stellar masses estimated from SED fitting is shown. The dotted line indicates equal stellar and dynamical masses. The average dynamic to stellar mass ratio is  $\sim 3$ , which is higher than, but consistent with, similar measurements at  $z \approx 1-3$  (ref. 28). Errors on the dynamical masses reflect the  $1\sigma$  measurement uncertainty in the size, and velocity dispersion of the sources, but the geometry is limited to  $\sin(i) = 0.45-1$  ('Derivation of physical parameters' in Methods). The errors on the stellar masses reflect the  $1\sigma$  range of masses allowed by the SED model fits including emission line strength as a free parameter.



#### Extended Data Figure 4 | The star-formation rate history of the Universe.

The global SFR history at  $z > 4$  derived from UV measurements<sup>2</sup> is shown for three different assumptions about the dust obscuration in the general galaxy population. Blue, no correction for dust ('direct UV measurements' in key); red, a dust correction assuming the Meurer relation<sup>29</sup> ('UV corrected with Meurer *et al.*'); green, a correction which linearly evolves in redshift between a Meurer relation at  $z \approx 4$ , the SMC-like relation at  $z \approx 5$ , and our measured value at  $z \approx 6$  ('UV corrected with evolving dust'). Note the evolving dust correction leads to a downward revision by 30–40% at  $z \approx 6$ . Redshift error bars reflect the binning of the data, errors in the SFR density reflect the  $1\sigma$  measurement error in the ultraviolet luminosity density.



Extended Data Table 1 | Measured source properties

| ID               | RA         | DEC      | Redshift (UV) <sup>  </sup> | Selection <sup>†</sup> | Redshift ([CII]) <sup>‡</sup> | Continuum Flux ( $\mu$ Jy) | Line Flux (Jy km/s) | $\sigma$ [CII] (km/s) <sup>‡</sup> |
|------------------|------------|----------|-----------------------------|------------------------|-------------------------------|----------------------------|---------------------|------------------------------------|
| HZ1              | 149.971869 | 2.118175 | 5.690                       | LAE/I-LBG              | 5.6885(2)                     | <30                        | 0.29 $\pm$ 0.04     | 72 $\pm$ 11                        |
| HZ2              | 150.517074 | 1.928902 | 5.670                       | LAE/I-LBG              | 5.6697(6)                     | <29                        | 0.42 $\pm$ 0.07     | 134 $\pm$ 25                       |
| HZ3              | 150.039297 | 2.337183 | 5.546                       | I-LBG                  | 5.5416(2)                     | <51                        | 0.56 $\pm$ 0.07     | 91 $\pm$ 12                        |
| HZ4              | 149.618835 | 2.051873 | 5.540                       | BBG                    | 5.5440(2)                     | 202 $\pm$ 47               | 1.14 $\pm$ 11       | 108 $\pm$ 11                       |
| HZ5 <sup>§</sup> | 150.214993 | 2.582653 | 5.310                       | V-LBG                  | No data                       | <32                        | <0.02               | No data                            |
| HZ5a             | 150.214667 | 2.583136 | No Data                     | [CII]                  | 5.3089(1)                     | <32                        | 0.18 $\pm$ 0.02     | 20 $\pm$ 3                         |
| HZ6              | 150.08960  | 2.586411 | 5.290                       | V-LBG                  | 5.2928(1)                     | 129 $\pm$ 36               | 1.80 $\pm$ 0.13     | 95 $\pm$ 7                         |
| HZ7              | 149.876986 | 2.134113 | 5.250                       | BBG                    | 5.2532(4)                     | <36                        | 0.71 $\pm$ 0.07     | 162 $\pm$ 18                       |
| HZ8              | 150.016896 | 2.626614 | 5.148                       | R-LBG                  | 5.1533(2)                     | <30                        | 0.34 $\pm$ 0.03     | 107 $\pm$ 9                        |
| HZ8W             | 150.016546 | 2.626731 | No Data                     | [CII]                  | 5.1532(1)                     | <30                        | 0.27 $\pm$ 0.03     | 63 $\pm$ 7                         |
| HZ9              | 149.965406 | 2.378379 | 5.548                       | R-LBG                  | 5.5410(1)                     | 516 $\pm$ 42               | 1.95 $\pm$ 0.07     | 152 $\pm$ 6                        |
| HZ10             | 150.247098 | 1.555426 | 5.659                       | LAE/I-LBG              | 5.6566(2)                     | 1261 $\pm$ 44              | 1.57 $\pm$ 0.09     | 127 $\pm$ 8                        |

RA, right ascension; DEC, declination.

\* This source is the same object reported as LBG-1 in ref. 25.

† LAE, Narrow band selected Ly $\alpha$  emitter; I-LBG, I band dropout Lyman-break galaxy; R-LBG, R band dropout Lyman-break galaxy; V-LBG, V band dropout Lyman-break galaxy; BBG, Balmer break galaxy, selected with photo-z and 4.5  $\mu$ m flux; [C II], [C II] detection in these data.

‡ Velocity dispersion, FWHM = 2.35 $\sigma$ .

§ Chandra detected quasar, shows broad lines in spectra.

|| Redshift error is  $\pm 4 \times 10^{-3}$ .

¶ The numbers in parentheses indicate the errors in the final digits of the redshifts.

Extended Data Table 2 | Gaussian fits to ALMA [C II] moment zero images

| ID   | Beam Major Axis (") | Beam Minor Axis (") | Beam PA (Degrees) | Source Major Axis (") | Source Minor Axis (") | Source PA (Degrees) |
|------|---------------------|---------------------|-------------------|-----------------------|-----------------------|---------------------|
| HZ1  | 0.82                | 0.51                | 64.2              | 0.93±0.08             | 0.57±0.03             | 69.7±4.2            |
| HZ2  | 0.91                | 0.5                 | 64.9              | 1.46±0.37             | 0.69±0.08             | 79.3±5.6            |
| HZ3  | 0.87                | 0.4                 | 63.0              | 1.02±0.26             | 0.59±0.1              | 94±12               |
| HZ4  | 0.90                | 0.48                | 59.7              | 1.22±0.15             | 0.61±0.05             | 43.1±4              |
| HZ5  | 0.61                | 0.49                | 65.1              | No Data               | No Data               | No Data             |
| HZ5a | 0.61                | 0.49                | 65.1              | 0.93±0.16             | 0.58±0.07             | 120±10              |
| HZ6  | 0.68                | 0.48                | 65.2              | 1.07±0.14             | 0.83±0.09             | 48±17               |
| HZ7  | 0.71                | 0.49                | 76.8              | 1.19±0.22             | 0.65±0.08             | 72.9±7.9            |
| HZ8  | 0.71                | 0.49                | 61.5              | 1.41±0.21             | 1.01±0.13             | 35±16               |
| HZ8W | 0.71                | 0.49                | 61.5              | 1.05±0.13             | 0.55±0.04             | 75.3±4.2            |
| HZ9  | 0.66                | 0.53                | 75.8              | 0.94±0.07             | 0.66±0.04             | 76.2±6.1            |
| HZ10 | 0.69                | 0.51                | 60.5              | 1.04±0.06             | 0.60±0.02             | 78.7±2.8            |

PA, position angle.  
\* Measured source sizes, not de-convolved.

Extended Data Table 3 | De-convolved sizes of [C II] moment zero images.

| ID   | Major Axis (") | Minor Axis (") | PA (Degrees) | Major Axis (kpc) | Minor Axis (kpc) |
|------|----------------|----------------|--------------|------------------|------------------|
| HZ1  | 0.45±0.24      | 0.19±0.1       | 84±59        | 2.6±1.4          | 1.1±0.6          |
| HZ2  | 1.17±0.38      | 0.41±0.38      | 86±21        | 6.9±2.2          | 2.4±2.2          |
| HZ3  | <1.02          | <0.59          | ±180         | <6.1             | <3.5             |
| HZ4  | 0.87±0.24      | 0.24±0.16      | 29±18        | 5.2±1.4          | 1.4±1.0          |
| HZ5a | <0.96          | <0.33          | ±180         | <5.8             | <2.0             |
| HZ6  | 0.85±0.21      | 0.65±0.20      | 35±60        | 5.2±1.3          | 4.0±1.22         |
| HZ7  | 0.96±0.30      | 0.42±0.17      | 72±23        | 5.9±1.8          | 2.6±1.0          |
| HZ8  | 1.24±0.28      | 0.84±0.23      | 28±48        | 7.7±1.7          | 5.2±1.4          |
| HZ8W | 0.79±0.19      | 0.19±0.14      | 81±12        | 4.9±1.2          | 1.2±0.9          |
| HZ9  | 0.66±0.09      | 0.39±0.07      | 76±14        | 3.9±0.5          | 2.3±0.4          |
| HZ10 | 0.80±0.09      | 0.28±0.08      | 85.3±5.9     | 4.7±0.5          | 1.7±0.5          |



Extended Data Table 4 | Gaussian fits to ALMA continuum images\*

| ID   | Beam Major Axis (") | Beam Minor Axis (") | Beam PA (Degrees) | Source Major Axis (") | Source Minor Axis (") | Source PA (Degrees) |
|------|---------------------|---------------------|-------------------|-----------------------|-----------------------|---------------------|
| HZ4  | 0.87                | 0.48                | 59.2              | 0.82±0.07             | 0.44±0.02             | 56.4±2.9            |
| HZ6  | 0.71                | 0.50                | 65.5              | 0.72±0.06             | 0.45±0.02             | 73.9±4.2            |
| HZ9  | 0.63                | 0.51                | 76.6              | 0.72±0.06             | 0.56±0.04             | 93±11               |
| HZ10 | 0.66                | 0.49                | 61.2              | 1.16±0.11             | 0.59±0.03             | 83.2±3.1            |

\*Only HZ10 is resolved in the continuum with a de-convolved size of major axis =  $0.98 \pm 0.13''$ , minor axis =  $0.27 \pm 0.12''$ , PA =  $87.5 \pm 5.5$ .

Extended Data Table 5 | Derived physical properties

| ID               | $L_{\text{sun}} \text{ UV}^{\dagger}$ | $L_{\text{sun}} \text{ IR}^{\ddagger, \S}$ | SFR<br>( $M_{\text{sun}} \text{ yr}^{-1}$ ) | $L_{\text{sun}} \text{ [CII]}$ | Beta $^{\S}$             | UV $r^{1/2}$<br>(kpc) | Stellar Mass<br>( $M_{\text{sun}}$ ) | Dynamical Mass<br>( $M_{\text{sun}}$ ) |
|------------------|---------------------------------------|--|---|--------------------------------|--------------------------|-----------------------|--------------------------------------|--|
| HZ1              | 11.15±0.05                            | <10.32                                     | $24^{+6}_{-3}$                              | 8.40±0.32                      | -1.41±0.13               | 1.53                  | 10.47±0.13                           | $9.8^{+1.6}_{-0.8}$                    |
| HZ2              | 11.16±0.03                            | <10.3                                      | $25^{+5}_{-2}$                              | 8.56±0.41                      | -0.99 $^{+0.41}_{-0.92}$ | 0.59                  | 10.23±0.15                           | $10.7^{+1.5}_{-1.0}$                   |
| HZ3              | 11.02±0.07                            | <10.53                                     | $18^{+8}_{-3}$                              | 8.67±0.28                      | -1.05±0.5                | 0.66                  | 10.23±0.35                           | <10.4 $^{+1.6}$                        |
| HZ4              | 11.26±0.05                            | 11.13±0.54                                 | $51^{+54}_{-18}$                            | 8.98±0.22                      | -1.20±0.36               | 0.72                  | 9.67±0.21                            | $10.4^{+1.6}_{-1.1}$                   |
| HZ5 <sup>#</sup> | 11.46±0.02                            | <10.3                                      | <3  | <7.2                           | -0.77 $^{+0.21}_{-0.47}$ | 0.37                  | No data                              | No data                                |
| HZ5a             | <10.3                                 | <10.3                                      | <3  | 8.15±0.27                      | No data                  | No data               | No data                              | <9.0 $^{+1.8}$                         |
| HZ6              | 11.33±0.03                            | 10.91±0.64                                 | $49^{+44}_{-12}$                            | 9.15±0.17                      | -1.13 $^{+0.22}_{-0.71}$ | 3.36                  | 10.17±0.15                           | $10.6^{+1.6}_{-0.8}$                   |
| HZ7              | 11.08±0.04                            | <10.35                                     | $21^{+5}_{-2}$                              | 8.74±0.24                      | -1.24±0.24               | 0.98                  | 9.86±0.21                            | $10.8^{+1.5}_{-1.0}$                   |
| HZ8              | 11.02±0.06 <sup>  </sup>              | <10.26                                     | $18^{+5}_{-2}$                              | 8.41±0.18                      | -1.42 $^{+0.15}_{-0.52}$ | 1.24                  | 9.77±0.15                            | $10.8^{+1.3}_{-0.9}$                   |
| HZ8W             | 10.57±0.15 <sup>  </sup>              | <10.26                                     | $6^{+5}_{-2}$                               | 8.31±0.23                      | No data                  | 1.49                  | No data                              | $9.9^{+1.6}_{-1.2}$                    |
| HZ9              | 10.93±0.05                            | 11.54±0.19                                 | $67^{+30}_{-20}$                            | 9.21±0.09                      | -0.85 $^{+0.22}_{-0.95}$ | 0.95                  | 9.86±0.23                            | $10.7^{+1.3}_{-1.1}$                   |
| HZ10             | 11.35±0.06                            | 11.94±0.08                                 | $169^{+32}_{-27}$                           | 9.13±0.13                      | <-0.6 <sup>§</sup>       | No data               | 10.39±0.17                           | $10.5^{+1.4}_{-1.2}$                   |

\* Monochromatic measured luminosity at rest frame 1,600 Å with no correction for obscuration.  $M_{1600} = -2.5L_{\text{UV}} + 5.814$ .

† ±0.3 systematic uncertainty in addition to quoted statistical errors.

‡ 3–1,100 μm integrated luminosity.

§ Systematic error is an additional ±0.3 dex based on simulations.

|| The total ultraviolet luminosity for both components was measured as  $11.15 \pm 0.04$  and split between the two components based on the ratio of their observed  $z'$  magnitudes.

# Based on the 5σ non-detection in H band.

# Known X-ray detected quasar.

\*\* The low SNR H band point is a significant outlier, error reflects range with that point included and rejected.

# A giant comet-like cloud of hydrogen escaping the warm Neptune-mass exoplanet GJ 436b

David Ehrenreich<sup>1</sup>, Vincent Bourrier<sup>1</sup>, Peter J. Wheatley<sup>2</sup>, Alain Lecavelier des Etangs<sup>3,4</sup>, Guillaume Hébrard<sup>3,4,5</sup>, Stéphane Udry<sup>1</sup>, Xavier Bonfils<sup>6,7</sup>, Xavier Delfosse<sup>6,7</sup>, Jean-Michel Désert<sup>8</sup>, David K. Sing<sup>9</sup> & Alfred Vidal-Madjar<sup>3,4</sup>

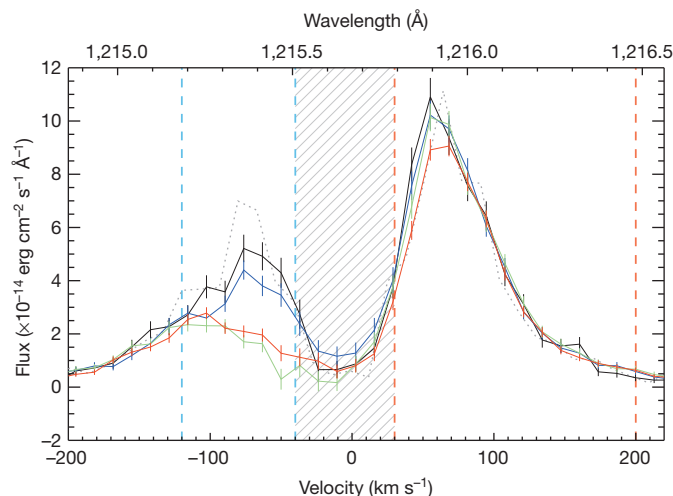
Exoplanets orbiting close to their parent stars may lose some fraction of their atmospheres because of the extreme irradiation<sup>1–6</sup>. Atmospheric mass loss primarily affects low-mass exoplanets, leading to the suggestion that hot rocky planets<sup>7–9</sup> might have begun as Neptune-like<sup>10–16</sup>, but subsequently lost all of their atmospheres; however, no confident measurements have hitherto been available. The signature of this loss could be observed in the ultraviolet spectrum, when the planet and its escaping atmosphere transit the star, giving rise to deeper and longer transit signatures than in the optical spectrum<sup>17</sup>. Here we report that in the ultraviolet the Neptune-mass exoplanet GJ 436b (also known as Gliese 436b) has transit depths of  $56.3 \pm 3.5\%$  ( $1\sigma$ ), far beyond the 0.69% optical transit depth. The ultraviolet transits repeatedly start about two hours before, and end more than three hours after the approximately one hour optical transit, which is substantially different from one previous claim<sup>6</sup> (based on an inaccurate ephemeris). We infer from this that the planet is surrounded and trailed by a large exospheric cloud composed mainly of hydrogen atoms. We estimate a mass-loss rate in the range of about  $10^8$ – $10^9$  grams per second, which is far too small to deplete the atmosphere of a Neptune-like planet in the lifetime of the parent star, but would have been much greater in the past.

Three transits of GJ 436b, which occur every 2.64 days, were observed on 7 December 2012 (ref. 6) (visit 1), 18 June 2013 (visit 2) and 23 June 2014 (visit 3) with the Space Telescope Imaging Spectrograph (STIS) on board the Hubble Space Telescope (HST). A stellar spectrum acquired using similar settings in January 2010 (ref. 17) (visit 0) was retrieved from the archive for comparison purposes. HST data in visits 2 and 3 were complemented with simultaneous Chandra X-ray observations. The HST data consist of time-tagged, far-ultraviolet spectra obtained with a grating dispersing light over the 1,195–1,248 Å domain, with a spectral resolution of  $\sim 20$  km s<sup>−1</sup> at 1,215.6 Å (the Lyman-α transition of atomic hydrogen, H I). Exposure times of 1,500 s to 2,900 s were used to observe the star for four successive HST orbits during each visit. Each HST orbit lasts for 96 min, during which GJ 436 is visible for 56 min before being occulted by the Earth, yielding 40 min gaps in the data.

The most prominent spectral feature is the H I Lyman-α emission (Fig. 1). Absorption in the blue wing of this line has been reported in other systems, during transits of hot Jupiters. This is interpreted by the presence of escaping hydrogen exospheres surrounding giant planets<sup>1,5,18–20</sup>. Tentative evidence that the Neptune-mass planet GJ 436b possesses such an extended atmosphere was drawn from visit 1 data despite the signal being observed after one optical transit<sup>6</sup>, raising questions about its planetary origin. Visits 2 and 3 were carried out to determine the signal nature.

We performed a careful analysis to check for the existence of instrumental systematics in the data and correct for them

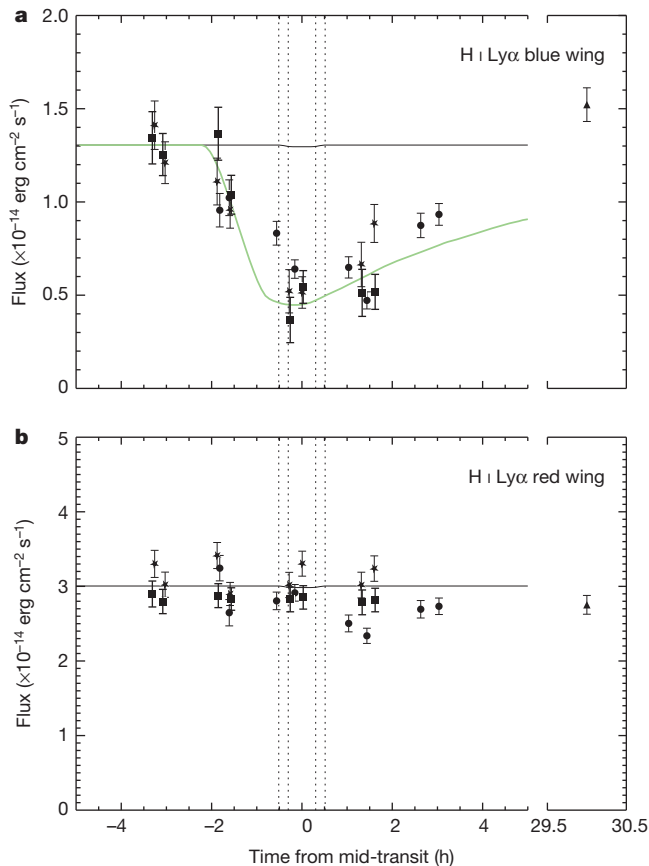
(see Methods). Large variations are detected over a part of the Lyman-α line at times corresponding to the optical transit, which cannot be explained by any known instrumental effects. The most notable absorption occurs in the blue wing of the line for radial velocities between  $-120$  km s<sup>−1</sup> to  $-40$  km s<sup>−1</sup>,  $\sim 2$  h before the optical mid-transit time ('pre-transit'), during the optical transit ('in-transit') and  $\sim 1$  h after the optical mid-transit time and beyond ('post-transit') in the three visits (Fig. 1 and Extended Data Fig. 1). In-transit, over half the stellar disk ( $56.2 \pm 3.6\%$ ,  $1\sigma$ ) is eclipsed (Fig. 2a). This is far deeper than any ultraviolet transits of hot Jupiters and significantly ( $\sim 9\sigma$ ) deeper than the  $22.9 \pm 3.9\%$  ( $1\sigma$ ) post-transit signal previously reported in visit 1 data<sup>6</sup>. The ultraviolet transit also starts much earlier ( $\sim 2.7$  h) than claimed previously; the difference is mainly due to our finding of a pre-transit absorption and updated transit ephemeris (see Methods). Visit 0 data<sup>17</sup> acquired  $\sim 3$  years before visit 1 and  $\sim 30$  h after transit show that the out-of-transit variability is small compared with the blue-shifted signature. This gives us confidence that the true out-of-transit baseline is measured in



**Figure 1 | Evolution of the hydrogen Lyman-α emission line of GJ 436.** The line has been averaged over out-of-transit (black), pre-transit (blue), in-transit (green) and post-transit (red) observations from individual spectra (see Extended Data Fig. 1). The  $1\sigma$  uncertainties have been propagated accordingly from the errors calculated by the STIS data reduction pipeline. The line profile from visit 0 is shown for comparison with a dotted grey line. The line core (hatched region) cannot be observed from Earth because of the interstellar medium absorption along the line of sight. Absorption signals are measured over the interval  $[-120, -40]$  km s<sup>−1</sup> (blue dashed lines) and compared to a control measure over the interval  $[+30, +200]$  km s<sup>−1</sup> (red dashed lines). The velocity scale is heliocentric.

<sup>1</sup>Observatoire de l'Université de Genève, 51 chemin des Maillettes, 1290 Versoix, Switzerland. <sup>2</sup>Department of Physics, University of Warwick, Coventry CV4 7AL, UK. <sup>3</sup>CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis boulevard Arago, 75014 Paris, France. <sup>4</sup>Sorbonne Universités, UPMC Univ. Paris 6, UMR 7095, Institut d'Astrophysique de Paris, 98 bis boulevard Arago, 75014 Paris, France. <sup>5</sup>Observatoire de Haute-Provence, CNRS & OAMP, 04870 Saint-Michel-l'Observatoire, France. <sup>6</sup>Univ. Grenoble Alpes, IPAG, F-38000 Grenoble, France. <sup>7</sup>CNRS, IPAG, F-38000 Grenoble, France. <sup>8</sup>CASA, Department of Astrophysical & Planetary Sciences, University of Colorado, 389-UCB, Boulder, Colorado 80309, USA. <sup>9</sup>Astrophysics Group, School of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL, UK.



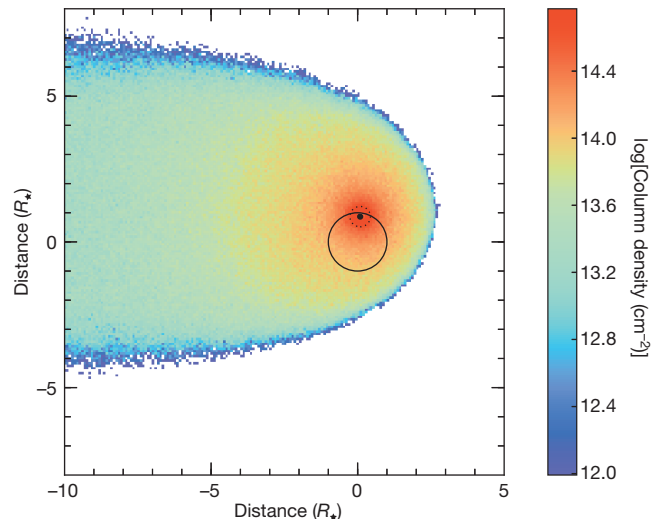


**Figure 2 | Lyman- $\alpha$  transit light curves of GJ 436b.** **a, b,** Data are from visit 1 (circles), visit 2 (stars), visit 3 (squares) and visit 0 (triangles). All uncertainties are  $1\sigma$ . **a,** The Lyman- $\alpha$  ( $\text{Ly}\alpha$ ) line is integrated over  $[-120, -40] \text{ km s}^{-1}$  and shows mean absorption signals with respect to the out-of-transit flux of  $17.6 \pm 5.2\%$  (pre-transit),  $56.2 \pm 3.6\%$  (in-transit) and  $47.2 \pm 4.1\%$  (post-transit). **b,** The line is integrated over  $[+30, +200] \text{ km s}^{-1}$  and shows no notable absorption signals:  $0.7 \pm 3.6\%$  (pre-transit),  $1.7 \pm 3.5\%$  (in-transit) and  $8.0 \pm 3.1\%$  (post-transit). With a depth of  $0.69\%$ , the optical transit (thin black lines in **a** and **b**) is barely seen at this scale between its contact points (dotted lines in **a** and **b**). A synthetic light curve (green) calculated from the three-dimensional numerical simulation<sup>20</sup> is overplotted on the data in **a**.

visits 2 and 3, whereas it is missing in visit 1. By contrast, the flux remains stable over the whole red-shifted wing of the line (Fig. 2b). The decrease of the red-wing flux seen<sup>6</sup> during the post-transit phases of visit 1 is not reproduced during visits 2 and 3. The mean post-transit red-shifted signal is compatible with no detection at the  $3\sigma$  level.

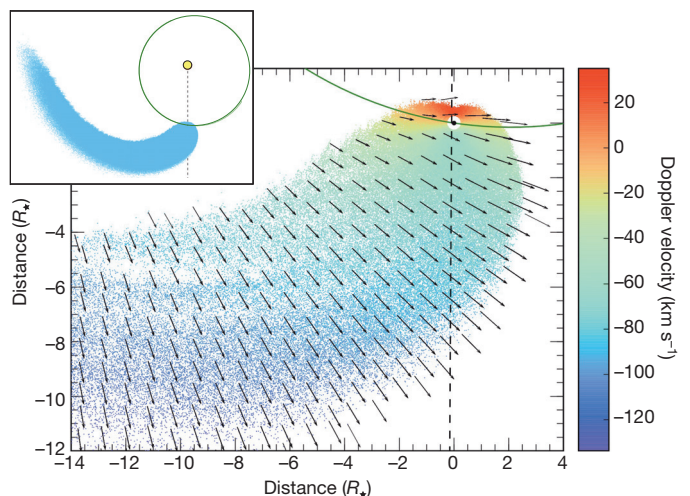
Our combined analysis of X-ray and ultraviolet data (see Methods) shows that stellar magnetic activity cannot explain the observed decrease at Lyman- $\alpha$ . We propose that the asymmetric absorption is caused by the passage of a huge hydrogen cloud, surrounding and trailing the planet (Fig. 3). The planetary atmosphere is an obvious source for this hydrogen. To produce this extinction signature, we estimate that an ellipsoidal, optically thick cloud of neutral hydrogen should have a projected extension in the plane of the sky of  $\sim 12$  stellar radii ( $R_{\star} \approx 0.44 R_{\odot}$ ) along the orbital path of the planet and  $\sim 2.5 R_{\star}$  in the cross direction, well beyond the planet Roche lobe radius ( $0.37 R_{\star}$ ). Since GJ 436b grazes the stellar disk during transit, we surmise that a central transit would have totally eclipsed the star. This could happen in the case of other red dwarfs exhibiting central transits from planets similar to GJ 436b. Future ultraviolet observations of systems similar to GJ 436 could potentially reveal total Lyman- $\alpha$  eclipses.

The radial velocity interval of the absorption signal constrains the dynamics of the hydrogen atoms and the three-dimensional structure of the exospheric cloud. The whole velocity range is in excess of the



**Figure 3 | Particle simulation showing the comet-like exospheric cloud transiting the star, as seen from Earth.** GJ 436b is the small black dot represented at mid-transit at  $0.8521 R_{\star}$  (ref. 26) from the centre of the star, which is represented by the largest black circle. The dotted circle around the planet represents its equivalent Roche radius. The colour of simulation particles denotes the logarithm of the column density of the cloud. The transit of this simulated cloud gives rise to absorption over the blue wing of the Lyman- $\alpha$  line as shown spectrally in Extended Data Fig. 2 and by the synthetic light curve in Fig. 2a.

planet escape velocity ( $\sim 26 \text{ km s}^{-1}$  at the planet surface), consistent with gas escaping from the planet. The acceleration mechanism of hydrogen atoms escaping from highly irradiated hot Jupiters is debated: after escaping the planets with initial velocities dominated by the orbital velocity ( $\sim 100 \text{ km s}^{-1}$  for GJ 436b in the host star reference frame), atoms are submitted to the stellar radiation pressure, interact with the stellar wind and are eventually ionized by stellar extreme ultraviolet (EUV;  $10\text{--}91.2 \text{ nm}$ ) radiation. For strong lines such as Lyman- $\alpha$ , radiation pressure can overcome the stellar gravity, repelling the escaping atoms towards the observer and producing a blue-shifted signature. In one hot Jupiter (HD 189733b), the absorption



**Figure 4 | Polar view of three-dimensional simulation representing a slice of the comet-like cloud coplanar with the line of sight.** Hydrogen atom velocity and direction in the rest frame of the star are represented by arrows. Particles are colour-coded as a function of their projected velocities on the line of sight (the dashed vertical line). Inset, zoom out of this image to the full spatial extent of the exospheric cloud (in blue). The planet orbit is shown to scale with the green ellipse and the star is represented with the yellow circle.

observed at large blue-shifts is best explained by charge exchange interaction with the stellar wind, creating energetic neutral atoms with large blue-shifted radial velocities<sup>19–22</sup>. In other cases<sup>1,18,23</sup>, radiation pressure alone explains the observed radial velocities of the escaping gas.

We ran a three-dimensional numerical simulation of atmospheric escape<sup>20</sup> to understand the origin of the absorption signature observed at GJ 436b. The line profile corrected from interstellar absorption<sup>17,24</sup> is used to calculate the stellar radiation pressure on hydrogen atoms. These are released isotropically from the Roche lobe limit of GJ 436b. The calculation takes the orbital eccentricity of GJ 436b ( $e = 0.15$ ) into account. The main parameters of the atmospheric escape model are the mass-loss rate  $\dot{m}$  of hydrogen lost by the planet and the photo-ionization rate  $\alpha_{\text{EUV}}$  of hydrogen atoms. The model computes the structure of the escaping gas cloud, its radial velocity absorption signature, and the integrated light curve. Adjusting for the spectra at different phases (Extended Data Fig. 2), we determine that a family of models, with parameters in the  $1\sigma$  ranges of  $\dot{m} \approx 10^8\text{--}10^9 \text{ g s}^{-1}$  and  $\alpha_{\text{EUV}} \approx 8 \times 10^{-7}\text{--}3 \times 10^{-6} \text{ s}^{-1}$  (implying neutral atom lifetimes of  $\sim 4\text{--}18$  min at the distance of the planet), provide good fits to the data, yielding reduced  $\chi^2$  values close to unity for 544 degrees of freedom. A synthetic light curve is plotted in Fig. 2. The model correctly matches the  $\sim 2$  h early ultraviolet transit ingress observed with respect to the optical transit, as well as the transit depth in the correct range of velocities. It provides a good match to the reanalysed visit 1 data. It furthermore predicts that the ultraviolet transit could last up to  $\sim 20$  h after the optical transit, due to the extended hydrogen tail of the exospheric cloud. More ultraviolet observations will be needed to verify this prediction.

According to our simulation, the stellar radiation pressure counterbalances  $\lesssim 70\%$  of the star's gravity pull on escaping atoms, which is much less than in hot Jupiter systems, in which radiation pressure takes over stellar gravity by factors of 3 to 5 (ref. 20). The low stellar radiation pressure at GJ 436b allows the formation of a large coma and tail of escaping atoms, co-moving with the planet although not gravitationally bounded to it (Fig. 4).

Atmospheric escape is involved in the possible loss of a whole population of irradiated exoplanets<sup>11,13,15,16</sup>. The average mass-loss rate of  $\sim 5 \times 10^8 \text{ g s}^{-1}$  at GJ 436b means that the planet loses  $\sim 0.1\%$  of its atmosphere per billion years (assuming it accounts for 10% of the planet mass, like Neptune). This rate requires  $\sim 1\%$  efficiency in the conversion of input X and EUV energy (estimated in Methods) into mass loss<sup>15</sup>. In the past, an M dwarf such as GJ 436 would have been more active and the planet could have received  $\lesssim 100$  times more X-ray and EUV irradiation over  $\sim 1 \text{ Gyr}$ <sup>25</sup>, resulting in a possible loss of  $\lesssim 10\%$  of its atmosphere during the first billion years. This planet thus stands on the edge of considerable mass loss, leading us to surmise that closer-in Neptunes could have evolved more dramatically because of atmospheric escape.

This  $16\sigma$  detection implies that large atmospheric signals from comet-like exospheres around moderately irradiated, low-mass planets could be retrieved in the ultraviolet. Over  $\sim 10,000$  systems similar to GJ 436 will be discovered by upcoming transit surveys carried out from the ground and from space—for example, with the K2, CHEOPS, TESS and PLATO missions.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 February; accepted 27 April 2015.

- Vidal-Madjar, A. *et al.* An extended upper atmosphere around the extrasolar planet HD209458b. *Nature* **422**, 143–146 (2003).
- Koskinen, T. T., Aylward, A. D. & Miller, S. A stability limit for the atmospheres of giant extrasolar planets. *Nature* **450**, 845–848 (2007).

- Linsky, J. L. *et al.* Observations of mass loss from the transiting exoplanet HD 209458b. *Astrophys. J.* **717**, 1291–1299 (2010).
- Fossati, L. *et al.* Metals in the exosphere of the highly irradiated planet WASP-12b. *Astrophys. J.* **714**, L222–L227 (2010).
- Lecavelier des Etangs, A. *et al.* Temporal variations in the evaporating atmosphere of the exoplanet HD 189733b. *Astron. Astrophys.* **543**, L4 (2012).
- Kulow, J. R., France, K., Linsky, J. & Loyd, R. O. P. Ly $\alpha$  transit spectroscopy and the neutral hydrogen tail of the hot Neptune GJ 436b. *Astrophys. J.* **786**, 132 (2014).
- Léger, A. *et al.* The extreme physical properties of the CoRoT-7b super-Earth. *Icarus* **213**, 1–11 (2011).
- Batalha, N. M. *et al.* Kepler's first rocky planet: Kepler-10b. *Astrophys. J.* **729**, 27 (2011).
- Pepe, F. *et al.* An Earth-sized planet with an Earth-like density. *Nature* **503**, 377–380 (2013).
- Lecavelier des Etangs, A., Vidal-Madjar, A., McConnell, J. C. & Hébrard, G. Atmospheric escape from hot Jupiters. *Astron. Astrophys.* **418**, L1–L4 (2004).
- Lecavelier des Etangs, A. A diagram to determine the evaporation status of extrasolar planets. *Astron. Astrophys.* **461**, 1185–1193 (2007).
- Hubbard, W. B., Hattori, M. F., Burrows, A. & Hubeny, I. A mass function constraint on extrasolar giant planet evaporation rates. *Astrophys. J.* **658**, L59–L62 (2007).
- Davis, T. A. & Wheatley, P. J. Evidence for a lost population of close-in exoplanets. *Mon. Not. R. Astron. Soc.* **396**, 1012–1017 (2009).
- Jackson, B. *et al.* The roles of tidal evolution and evaporative mass loss in the origin of CoRoT-7 b. *Mon. Not. R. Astron. Soc.* **407**, 910–922 (2010).
- Ehrenreich, D. & Désert, J.-M. Mass-loss rates for transiting exoplanets. *Astron. Astrophys.* **529**, A136 (2011).
- Lopez, E. D., Fortney, J. J. & Miller, N. How thermal evolution and mass-loss sculpt populations of super-Earths and sub-Neptunes: application to the Kepler-11 system and beyond. *Astrophys. J.* **761**, 59 (2012).
- Ehrenreich, D., Lecavelier des Etangs, A. & Delfosse, X. HST/STIS Lyman- $\alpha$  observations of the quiet M dwarf GJ 436. Predictions for the exospheric transit signature of the hot Neptune GJ 436b. *Astron. Astrophys.* **529**, A80 (2011).
- Ehrenreich, D. *et al.* Hint of a transiting extended atmosphere on 55 Cancri b. *Astron. Astrophys.* **547**, A18 (2012).
- Bourrier, V. *et al.* Atmospheric escape from HD 189733b observed in H I Lyman- $\alpha$ : detailed analysis of HST/STIS September 2011 observations. *Astron. Astrophys.* **551**, A63 (2013).
- Bourrier, V. & Lecavelier des Etangs, A. 3D model of hydrogen atmospheric escape from HD 209458b and HD 189733b: radiative blow-out and stellar wind interactions. *Astron. Astrophys.* **557**, A124 (2013).
- Holmström, M. *et al.* Energetic neutral atoms as the explanation for the high-velocity hydrogen around HD 209458b. *Nature* **451**, 970–972 (2008).
- Ekenbäck, A. *et al.* Energetic neutral atoms around HD 209458b: estimations of magnetospheric properties. *Astrophys. J.* **709**, 670–679 (2010).
- Bourrier, V., Lecavelier des Etangs, A. & Vidal-Madjar, A. Modeling magnesium escape from HD 209458b atmosphere. *Astron. Astrophys.* **565**, A105 (2014).
- France, K. *et al.* The ultraviolet radiation environment around M dwarf exoplanet host stars. *Astrophys. J.* **763**, 149 (2013).
- Ribas, I., Guinan, E. F., Güdel, M. & Audard, M. Evolution of the solar activity over time and effects on planetary atmospheres. I. High-energy irradiances (1–1700 Å). *Astrophys. J.* **622**, 680–694 (2005).
- Knutson, H. A. *et al.* A *Spitzer* transmission spectrum for the exoplanet GJ 436b, evidence for stellar variability, and constraints on dayside flux variations. *Astrophys. J.* **735**, 27 (2011).

**Acknowledgements** This work is based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with programmes #11817, #12034 and #12965. The scientific results reported in this article are based on observations made by the Chandra X-ray Observatory. This work was carried out in the framework of the National Centre for Competence in Research 'PlanetS' supported by the Swiss National Science Foundation (SNSF). D.E., V.B. and S.U. acknowledge the financial support of the SNSF. V.B., A.L.d.E., X.B. and X.D. acknowledge the support of CNES, the French Agence Nationale de la Recherche (ANR) under program ANR-12-BS05-0012 'Exo-Atmos', the Fondation Simone et Cino Del Duca, and the European Research Council (ERC) under ERC Grant Agreement no. 337591-ExTrA.

**Author Contributions** D.E. proposed and led the HST–Chandra joint observation programme, supervised data reduction and analysis, interpreted the results and wrote the paper. V.B. performed data reduction and analysis, and computer simulations to interpret the results. P.J.W. set up the Chandra X-ray observations, reduced, analysed and interpreted the X-ray data. A.L.d.E. co-designed the simulation programme with V.B. and provided computing resources to run the simulations. A.L.d.E. and G.H. contributed to the observation programme, data analysis and interpretation. S.U., X.B., X.D., J.-M.D., D.K.S. and A.V.-M. contributed to the observation programme and interpretation. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.E. ([david.ehrenreich@unige.ch](mailto:david.ehrenreich@unige.ch)).

## METHODS

**Observation log.** This work is mainly based on observations acquired in the framework of the joint HST–Chandra programme “Properties and dynamics of the upper atmosphere of the hot-Neptune GJ 436b” (GO number 12965; principal investigator: D.E.). The programme consisted of three HST/STIS transit observations, with contemporaneous coverage of the pre-transit phases with Chandra/ACIS-S. During the first joint HST–Chandra visit, initially scheduled on 16 February 2013, we only obtained Chandra data as a last-minute mission scheduling change prevented the HST observation. It was rescheduled, together with a new Chandra observation, on 18 April 2013. However, no useful HST data could again be obtained on this date because of a target pointing problem. Meanwhile, the simultaneous Chandra observation was correctly performed. The remaining two HST–Chandra visits were smoothly carried out on 18 June 2013 and 23 June 2014.

We complement our new data sets with archive HST/STIS observations of GJ 436: (1) a short observation of the star obtained  $\sim 30$  h after the planetary transit on 5 January 2010 (HST programme GO/DD number 11817 entitled “Detecting the upper atmosphere of a transiting hot Neptune: a feasibility study”; principal investigator: D.E.<sup>17</sup>); and (2) a transit observation obtained on 7 December 2012 (HST programme GTO number 12034 entitled “Brown dwarf activity part 2”; principal investigator: J. C. Green). The 7 December 2012 transit observation was previously published<sup>6</sup> and is discussed later. All HST and Chandra data are listed in Extended Data Table 1.

Similarly, we add archive X-ray data of GJ 436 to our new Chandra observations. These are (1) ROSAT All-Sky Survey data<sup>27</sup> and (2) XMM-Newton data<sup>28</sup> (observation identification number 556560101; principal investigator: P.J.W.).

**Data analysis and correction of systematics.** Because Lyman- $\alpha$  observations are photon-starved, the exposures are as long as the visibility of the target by HST allow. This yields spectra with exposure times of  $\sim 1,500$  s to  $\sim 2,900$  s, separated by  $\sim 40$  min gaps. The data are acquired in time-tagged mode, which means that the arrival time of each photon hitting the detector is recorded by the far-ultraviolet multi-anode microchannel array (FUV-MAMA) of STIS. It is thus possible to split the exposures into several sub-exposures. We adjusted the sub-exposure time to obtain a good compromise between time sampling and signal-to-noise ratio. (The time sampling is important to assess systematic effects, as described later.) During visit 1, the total exposure time was 1,515 s (orbit 1) and 2,905 s (orbits 2 to 4). The first orbit in a visit always has less science time available before of the target acquisition. We used  $4 \times 380$  s sub-exposures (orbit 1) and  $5 \times 581$  sub-exposures (orbits 2 to 4). During visits 2 and 3, the total exposure time was 1,670 s (orbit 1) and 2,071 s (orbits 2 to 4). This was subdivided into  $4 \times 417$  sub-exposures (orbit 1) and  $4 \times 518$  s sub-exposures (orbits 2 to 4), respectively. Extended Data Figs 3–5 show these time series, which are binned by a factor of 2 in Fig. 2 in the main text.

Variation of the temperatures in the optical telescope assembly and on the focal plan of HST during its revolution around the Earth is surmised to cause small changes of focus (for example, due to variation in the distance between the two M1 and M2 mirrors). For a given line of sight (hence for a given illumination pattern of the telescope), these variations depend on the HST orbital phase. They are consequently reproducible over several consecutive HST orbits, which means that their effect on the data, when present, can be corrected. This effect, known as the ‘telescope breathing’ effect, has been observed and corrected in STIS/CCD optical data<sup>29</sup> as well as in STIS/FUV-MAMA Lyman- $\alpha$  data<sup>18</sup>.

We use the red part of the Lyman- $\alpha$  line, within the velocity range  $[-20, +250]$  km s<sup>-1</sup>, as a probe for the systematic effects in general, and for the breathing effect in particular. In fact, the red peak and wing of the line are the only part of the spectrum unaffected (a priori) by the planet absorption and where there is enough flux to detect time-correlated effects. For each HST visit, we generate a raw light curve by integrating the Lyman- $\alpha$  flux over the velocity range given earlier, using the time-tag split exposures. These light curves are plotted in Extended Data Figs 3a, 4a and 5a, for visits 1, 2 and 3, respectively. If present in the data, the breathing effect would appear as a correlation in the light curves phase-folded over the HST orbital period (96 min). The phase-folded light curves for the three visits are shown in Extended Data Figs 3b, 4b and 5b. We then attempt to fit a polynomial to each phase-folded light curve and examine the dispersion of the residuals in the observed minus calculated data. When a significant correlation is found (as in visits 2 and 3), a phase-dependent correcting coefficient is applied to the whole spectrum.

For visit 1 data, we do not notice any significant correlation and find that a constant (0th-order polynomial) provides a better fit to the data ( $\chi^2 = 8.92$  for 16 degrees of freedom with a Bayesian information criterion (BIC) of 17.04) than any higher-order polynomials. Consequently, we do not apply any breathing correction for this visit. For visit 2 data, we do find an improvement (over a constant) in the fit of the phase-folded light curve with straight lines ( $\chi^2 = 0.88$  for 14 degrees of

freedom, BIC = 7.12) and apply this correction to the raw data. For visit 3 data, we find a linear correlation in the phase-folded light curve. We correct the raw data with the best-fit straight lines ( $\chi^2 = 2.23$  for 14 degrees of freedom, BIC = 12.7). The light curves corrected for breathing effect are shown in Extended Data Figs 4c and 5c for visits 2 and 3, respectively.

After having corrected for HST orbital phase-dependent effects, we now look at instrumental effects occurring on the timescale of a full HST visit (four HST orbits). Such effects typically include ‘ramps’ linked with long-term thermal relaxation of HST instruments. Our starting point is the Lyman- $\alpha$  red-wing light curves corrected for the breathing effect. Visit 1 still does not show any time correlation. In contrast, the visit 2 and 3 data show hints for a flux linearly decreasing with time. We chose to correct for this trend, by fitting a straight line to the data. This procedure yields a time-dependent correction coefficient for each visit. The final light curves corrected for (divided by) these trends are shown in Extended Data Figs 4d and 5d. The correction coefficients are applied to all spectra in visits 2 and 3. The data analysis and interpretation presented in the main text is based on these corrected spectra. These spectra are plotted for each visit in Extended Data Fig. 1. They clearly show that the signal is reproduced at each epoch, which strengthens our detection. In the main text, Fig. 1 presents these spectra averaged over visits 1, 2 and 3.

The amplitude of the correction of systematic effects performed as described earlier, either at the timescale of one HST orbit (breathing) or one HST visit (normalization), is small with respect to the absorption signal detected on the blue part of the Lyman- $\alpha$  line. The blue absorption would show up around the transit epoch almost unchanged, even in the absence of both corrections described earlier. Finally, the absorption signal by the exospheric cloud is so large that it can be measured directly in physical flux units. Our detection is thus independent of any normalization of the transit light curve.

**Determination of the velocity interval of the absorption signature.** In each individual spectrum of a given HST visit, we searched for the most significant (best signal-to-noise ratio) absorption signature. We scanned the velocity range  $[-250, +250]$  km s<sup>-1</sup>, excluding the low signal-to-noise-ratio region between  $-30$  km s<sup>-1</sup> and  $+20$  km s<sup>-1</sup>, for absorption signals extending over at least three contiguous elements of resolution. The signals are estimated from the difference between the spectrum and the reference out-of-transit spectrum for each visit. For visits 2 and 3, we simply used the data collected during the first HST orbits. We indifferently searched for absorption or emission signals, but found only significant ( $>3\sigma$ ) absorption features. We found that the range  $[-120, -40]$  km s<sup>-1</sup> captures most of the absorption features detected during these two visits. For visit 1, where the first orbit is already absorbed (since it is in the pre-transit phase), we could not rely on contemporaneous out-of-transit data for reference. We thus adopted the range used for visits 2 and 3.

On the basis of the assumption that observations obtained during the first HST orbit of visit 1 are not absorbed, ref. 6 measured a  $22.9 \pm 3.9\%$  ( $1\sigma$ ) absorption signal, with respect to this first HST orbit, over the range of  $[-215, -17]$  km s<sup>-1</sup>. We verified that we would obtain a similar result ( $\sim 24\%$ ) by making the same assumptions as in ref. 6. However, as we show here, the first exposures obtained during visit 1 already exhibit strong absorption by the exospheric cloud, and choosing the out-of-transit references of visits 2 and 3 leads to a much deeper absorption signal being measured.

**Transit ephemeris.** Several ephemerides of GJ 436b transit have been published in the literature<sup>6,26,30–32</sup> and we list them in Extended Data Table 2. Visit 1 transit observation has been published previously<sup>6</sup>, which reports the use of the following ephemeris values:  $P = 2.643850(90)$  days and  $T_0 = 2,454,279.436714(15)$ , for the period and mid-transit time, respectively. They attribute the period value to a previous study<sup>30</sup>, although the values eventually reported in these two studies<sup>6,30</sup> are different, as can be seen in Extended Data Table 2.

We recalculate the transit phases of the visit 1 observation using one of the most recent available ephemerides<sup>26</sup>. After taking into account the  $\sim 6$  min offset between the HST time values expressed in modified Julian days (MJD) and the barycentric Julian days (BJD<sub>T</sub>), we obtained the phase values for the four HST orbits of the visit 1 observation ( $-1.74$  h,  $-0.35$  h,  $+1.25$  h and  $+2.85$  h). The obtained values are all  $\sim 42$  min in advance of the phase values reported by ref. 6 ( $-1.05$  h,  $+0.35$  h,  $+1.95$  h and  $+3.54$  h). Trying to understand where this substantial difference came from, we realized that we would obtain similar transit phases as ref. 6, offset by  $\sim 6$  min, when using the  $T_0$  value of ref. 30 and the period given by ref. 6, which does not correspond to any value previously reported in the literature. The remaining  $\sim 6$  min difference presumably comes from our accounting for the delay between BJD and MJD<sub>UTC</sub>.

**New interpretation of a previous analysis of the visit 1 data.** The 42 min offset has two important consequences with respect to the visit 1 data interpretation by ref. 6. These authors indeed detected an absorption signature starting at the very end of the optical transit, with a maximum ultraviolet absorption depth of  $22.9 \pm$



3.9% ( $1\sigma$ ) occurring post-transit ( $\sim 2$  h after the end of the optical transit), according to their ephemeris. Shifting their data points 42 min earlier brings back the start of their absorption signal in-transit. The new phase of the first HST orbit in visit 1 is now  $-1.74$  h before the optical mid-transit time, which corresponds to the pre-transit phases in visits 2 and 3. Since we observe a significant ( $>3\sigma$ ) absorption signal of  $17.7 \pm 5.2\%$  in these visits, it means that the first HST orbit from visit 1, that ref. 6 used as an out-of-transit reference, is already partially absorbed by the exospheric cloud. In fact, the flux measured during the first HST orbits of visits 2 and 3 is significantly ( $>3\sigma$ ) higher than the flux measured in the second orbits of these visits (by the value of  $17.7 \pm 5.2\%$  quoted earlier), and in good agreement with the flux measured  $\sim 30$  h after transit, several years before in visit 0 (ref. 17), as can be seen in Fig. 2. The absorbed flux measured in the second HST orbits of visits 2 and 3 is a close match to the flux measured during the first HST orbit of visit 1. Using updated and more accurate transit ephemerides thus allows us to conciliate all HST observations within a common picture, quite different from the one painted by ref. 6. Not only does this substantially change the timing of the ultraviolet transit, but it also changes the amplitude of the measured absorption by a factor of  $\sim 2.5$ . This is critical for understanding the hydrogen atom dynamics and the exospheric cloud structure.

**X-ray observations and EUV flux of GJ 436.** Magnetic activity could induce variations in the stellar emission. We use the unabsorbed red part of the line to estimate that the intrinsic stellar variability at GJ 436 should not exceed 5% to 11% of the Lyman- $\alpha$  flux during one visit. Our Chandra X-ray data cover pre-transit phases at four epochs, two of which during HST visits 2 and 3, and we use them to monitor stellar activity. We detect a faint and soft X-ray source at the coordinates of GJ 436 in each of the four Chandra observations, with count rates of  $2.0 \pm 0.3$ ,  $3.1 \pm 0.4$ ,  $2.4 \pm 0.4$  and  $2.6 \pm 0.4$   $\text{ks}^{-1}$  for the February, April, June 2013 and June 2014 visits respectively. We extracted light curves (Extended Data Fig. 5) and spectra for each of these visits, as well as a combined spectrum for all four visits (Extended Data Fig. 6).

We also reanalysed the XMM-Newton EPIC-pn spectrum of GJ 436 (ref. 28) (Extended Data Fig. 6) and found a higher X-ray flux and luminosity than previously reported. We find an acceptable fit to the XMM-Newton spectrum with a two temperature thermal plasma model (APEC) with fitted temperatures of 0.09 and 0.38 keV and abundances of 0.18 Solar (fixed to the value listed by ref. 28). We measure a flux of  $4.6 \times 10^{-14}$   $\text{erg s}^{-1} \text{cm}^{-2}$  in the band previously used<sup>28</sup> (0.12–2.48 keV or 0.5–10 nm), giving a luminosity in this band of  $5.7 \times 10^{26}$   $\text{erg s}^{-1}$  (ref. 28 values were  $7.4 \times 10^{-15}$   $\text{erg s}^{-1} \text{cm}^{-2}$  and  $9.1 \times 10^{25}$   $\text{erg s}^{-1}$ , respectively). Our higher values for flux and luminosity are in much closer agreement with the measurement made from the ROSAT All-Sky Survey data for GJ 436 (ref. 27). We also inspected the ROSAT data and confirmed the marginal detection of the source with approximately eight photons, all below 0.3 keV.

We fitted the combined spectrum from our four Chandra observations simultaneously with the XMM-Newton EPIC-pn spectrum (Extended Data Fig. 6). The Chandra and XMM-Newton spectra match well and we find an acceptable simultaneous fit with the two-temperature model described earlier. The temperatures and relative normalization of the two components were left free but forced to take the same values for the two data sets. We find temperatures of 0.097 keV and 0.72 keV, or approximately 1 MK and 8 MK. Only the fluxes were allowed to vary between the XMM-Newton and Chandra spectra, and this combined fit can be seen in Extended Data Fig. 6 (XMM-Newton PN is black, Chandra ACIS-S is red). In the overlapping energy range (0.234–2.0 keV) X-ray fluxes are  $1.84 \times 10^{-14}$   $\text{erg s}^{-1} \text{cm}^{-2}$  for XMM-Newton (December 2008) and  $1.97 \times 10^{-14}$   $\text{erg s}^{-1} \text{cm}^{-2}$  for Chandra (averaged across the four observations in 2013 and 2014).

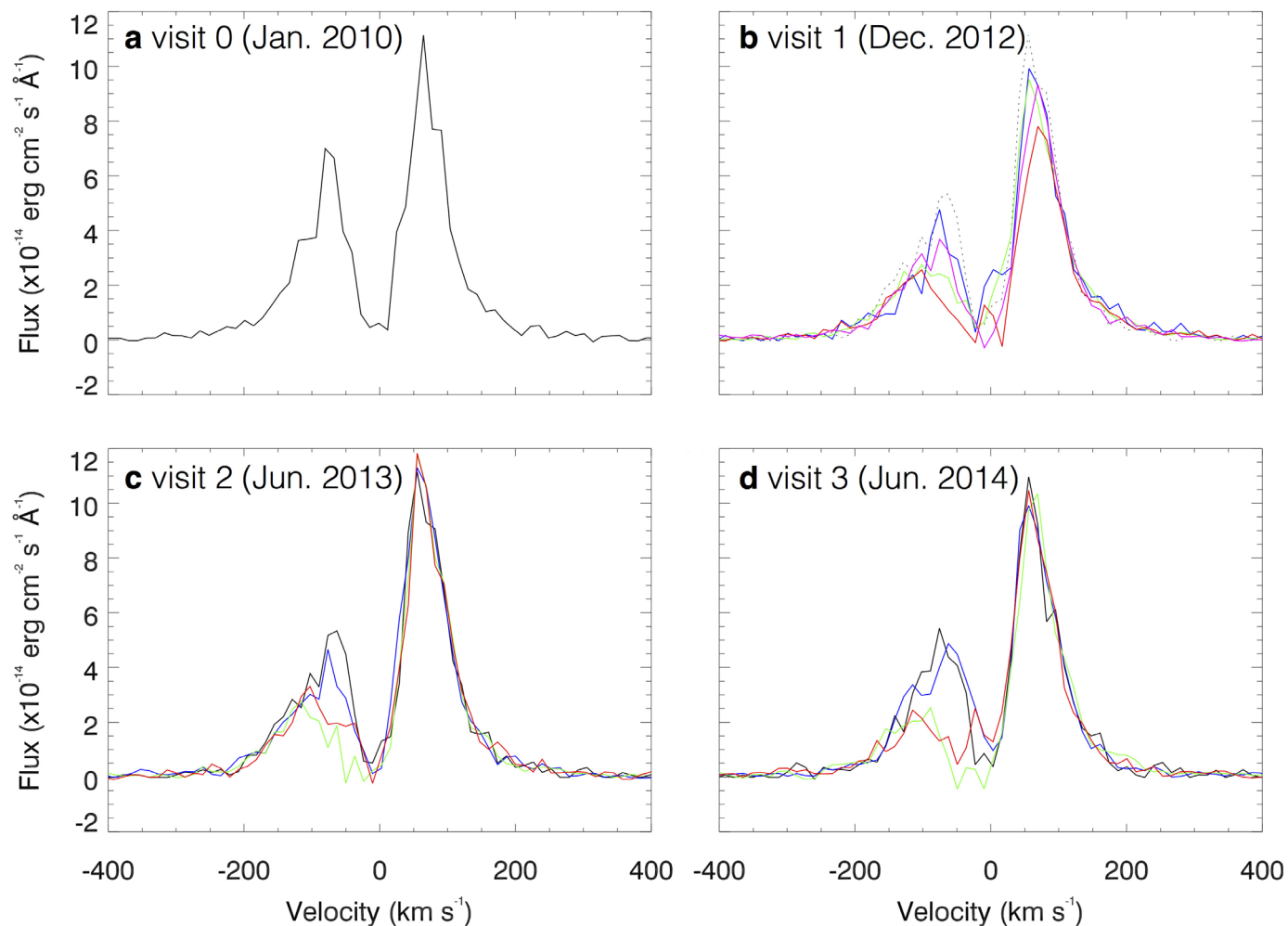
We also fitted the four Chandra spectra individually, fixing the two temperatures to those derived from the combined fit to the XMM-Newton and Chandra data. This allowed us to determine the emission measures of the two temperature components as a function of time, which are plotted in Extended Data Fig. 7. We find that the higher temperature component, which dominates the spectrum above 0.5 keV, is more variable than the lower temperature component (dominating below 0.5 keV), as shown by Extended Data Fig. 8. The variability of the higher temperature component accounts for most of the count rate variability in the Chandra observations and is presumably caused by stellar flares. The relatively stable lower temperature component is a useful proxy for the unobserved EUV emission that dominates the high-energy irradiation of the planet and probably drives the mass loss. The final two Chandra observations are contemporaneous with our HST visits 2 and 3, and while the emission measure of the high temperature component increases by a factor 2.3 between these observations, the emission measure of the low temperature component increases by only 7%. This helps to explain the very similar ultraviolet absorption depths and implied mass loss rates seen in HST visits 2 and 3. It can also be seen in Extended Data Fig. 8 that the emission measure of the lower temperature component was much lower during the first two Chandra observations in February and April 2013. HST visits at these times might have observed lower mass loss rates from the planet.

We use the X-ray flux measured in our third Chandra observation (contemporaneous with HST visit 2) to estimate the EUV flux irradiating the planet. The fitted model implies a flux in the 0.12–2.48 keV band (0.5–10 nm) of  $4.7 \times 10^{-14}$   $\text{erg s}^{-1} \text{cm}^{-2}$  (consistent with the flux at the time of the XMM-Newton observation). Converting to a flux at the surface of the star and comparing with the relation in figure 2 of ref. 33 leads to an estimate of the EUV flux at a reference separation of 1 au of  $1.0$   $\text{erg s}^{-1} \text{cm}^{-2}$  (12.4–91.2 nm), which is a factor of 5 higher than the X-ray flux at 1 au of  $0.21$   $\text{erg s}^{-1} \text{cm}^{-2}$  (0.5–10 nm).

This X-ray-derived EUV flux compares favourably with an estimate based on the strength of the Lyman- $\alpha$  line. Taking the total reconstructed line flux of  $3.5 \times 10^{-13}$   $\text{erg s}^{-1} \text{cm}^{-2}$  (ref. 24) and using the model from table 6 of ref. 34 gives an estimated EUV line flux at 1 au of  $1.06$   $\text{erg s}^{-1} \text{cm}^{-2}$  (10–91.2 nm). The consistency of the EUV flux estimates based on our X-ray and Lyman- $\alpha$  observations gives confidence that the high energy irradiation of the planet is well constrained despite being unobservable directly.

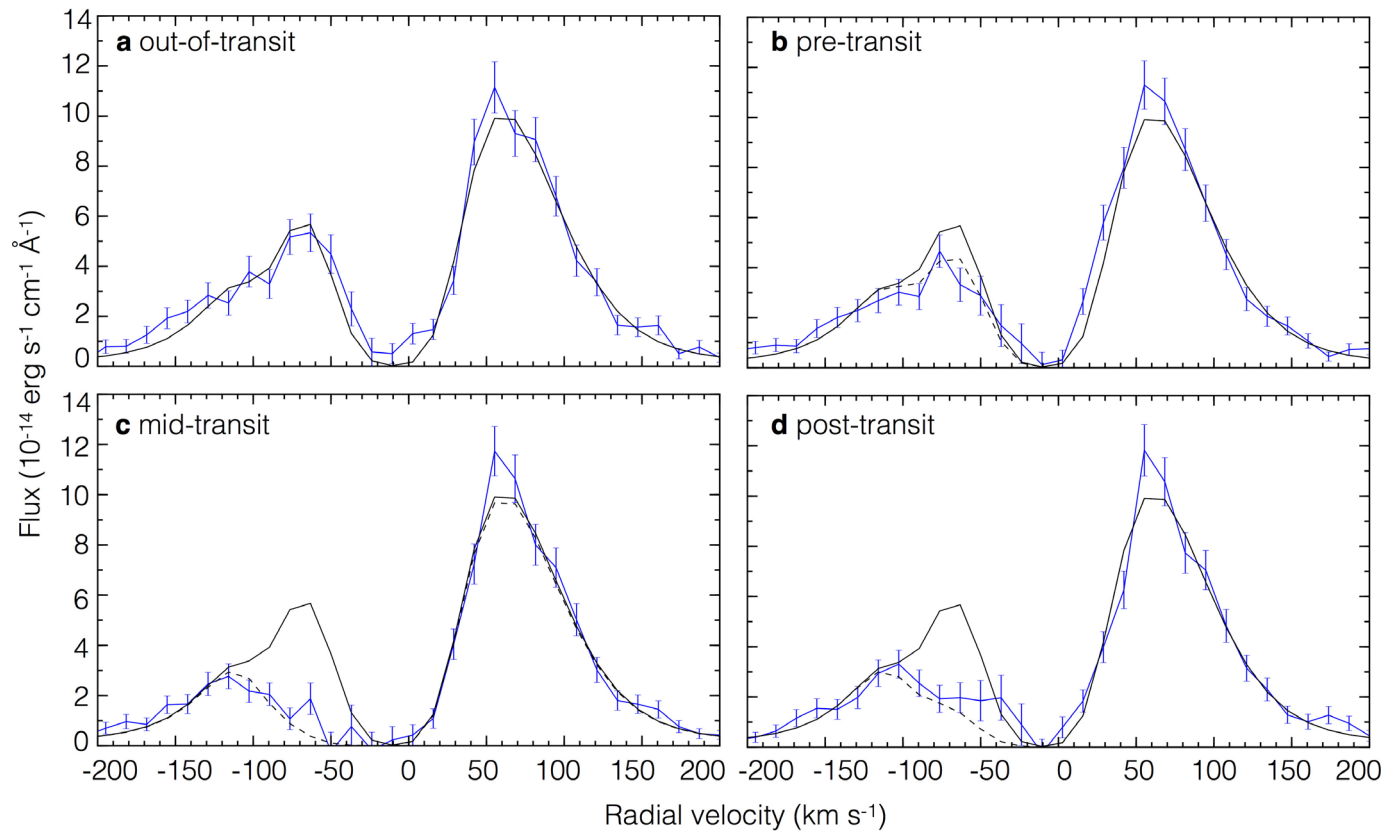
**Code availability.** We have opted not to make the code used to produce the numerical simulation available.

27. Poppenhaeger, K., Roßbrade, J. & Schmitt, J. H. M. M. Coronal properties of planet-bearing stars. *Astron. Astrophys.* **515**, A98 (2010).
28. Sanz-Forcada, J. *et al.* Estimation of the XUV radiation onto close planets and their evaporation. *Astron. Astrophys.* **532**, A6 (2011).
29. Demory, B.-O. *et al.* Hubble Space Telescope search for the transit of the Earth-mass exoplanet  $\alpha$  Centauri Bb. *Mon. Not. R. Astron. Soc.* **450**, 2043–2051 (2015).
30. Pont, F., Gilliland, R. L., Knutson, H., Holman, M. & Charbonneau, D. Transit infrared spectroscopy of the hot Neptune around GJ 436 with the Hubble Space Telescope. *Mon. Not. R. Astron. Soc.* **393**, L6–L10 (2009).
31. Knutson, H. A., Benneke, B., Deming, D. & Homeier, D. A featureless transmission spectrum for the Neptune-mass exoplanet GJ 436b. *Nature* **505**, 66–68 (2014).
32. Lanotte, A. *et al.* A global analysis of Spitzer and new HARPS data confirms the loneliness and metal-richness of GJ 436b. *Astron. Astrophys.* **572**, A73 (2014).
33. Chadney, J. M., Galand, M., Unruh, Y. C., Koskinen, T. T. & Sanz-Forcada, J. XUV-driven mass loss from extrasolar giant planets orbiting active stars. *Icarus* **250**, 357–367 (2015).
34. Linsky, J. L., Fontenla, J. & France, K. The intrinsic extreme ultraviolet fluxes of F5V to M5V Stars. *Astrophys. J.* **780**, 61 (2014).



**Extended Data Figure 1 | Time evolution of GJ 436 Lyman- $\alpha$  line during each HST visit.** There is one spectrum per HST orbit. Colours indicate the phase with respect to the optical transit: out-of-transit (black), pre-transit

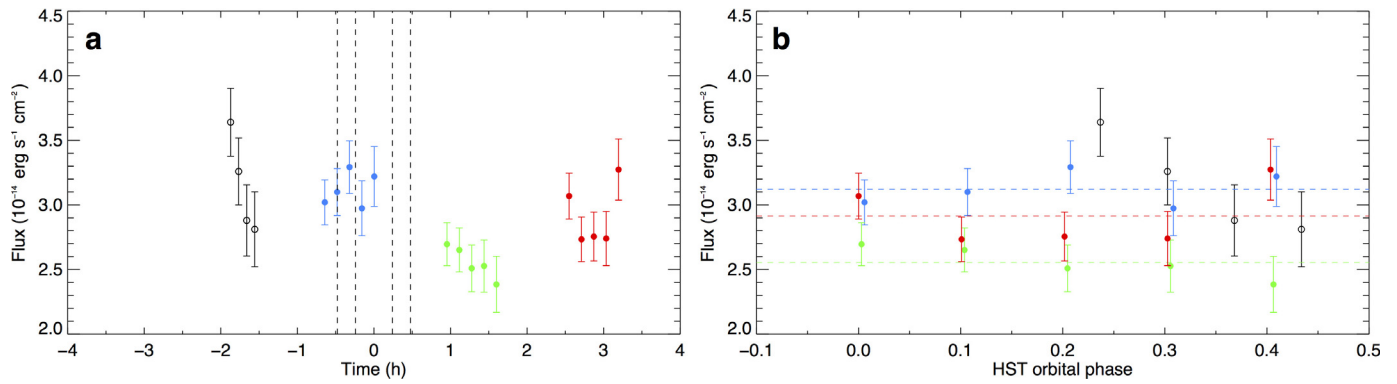
(blue), in-transit (green) and post-transit (red and magenta for the last HST orbit in visit 1). For visit 1 there is no out-of-transit spectrum, hence we over-plotted the out-of-transit spectrum from visit 2 (dotted line).



**Extended Data Figure 2 | Evolution of the Lyman- $\alpha$  line of GJ 436 reproduced with the numerical simulation.** **a–d**, The panels show the stellar emission line at different transit phases, here for visit 2 (all error bars are  $1\sigma$ ). **a**, The out-of-transit reference line (black curve) is the calculated profile best-fitting the observed out-of-transit line profile (blue curve), after taking into

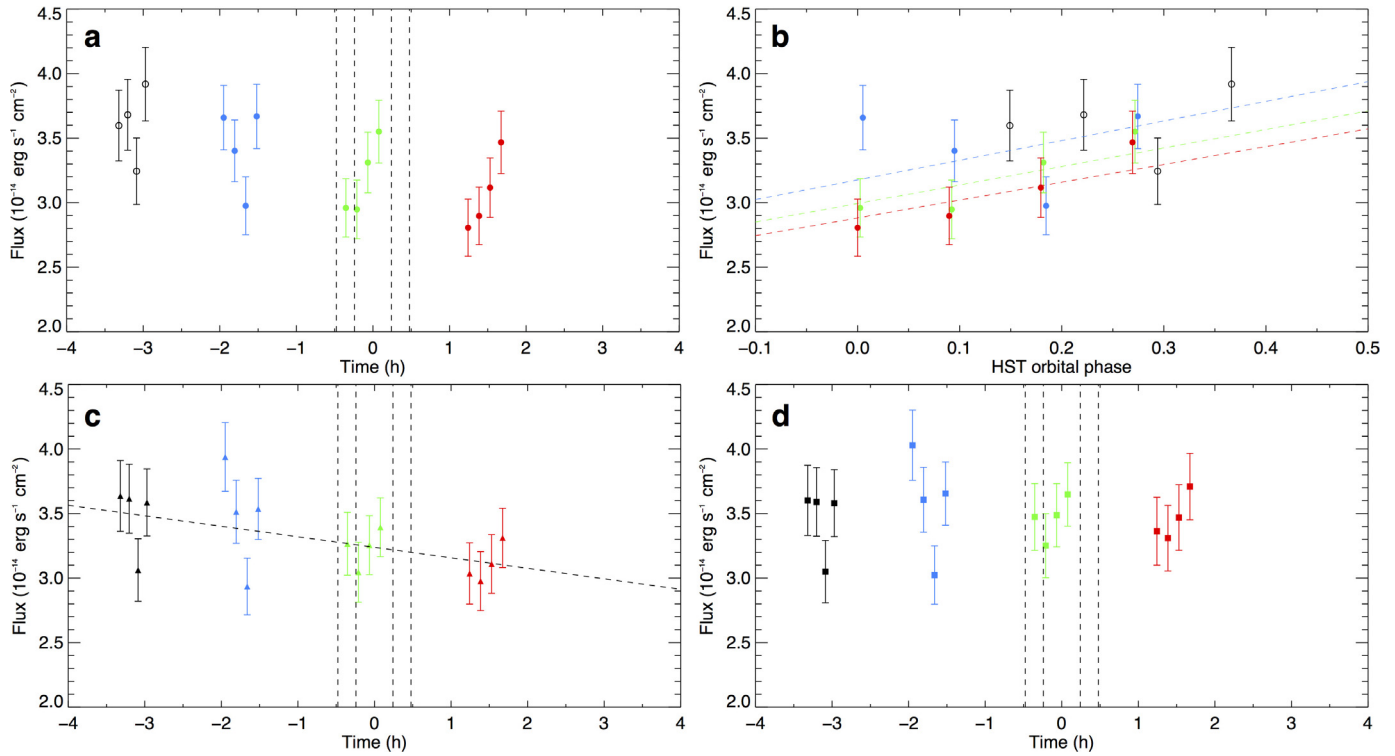
account the interstellar medium absorption and convolution by the instrumental line spread function. **b–d**, This theoretical profile is compared to observations (blue curves): the numerical simulation computes absorption in the theoretical profile and adjusts absorbed line profiles (dashed lines) as a function of time to the observations.





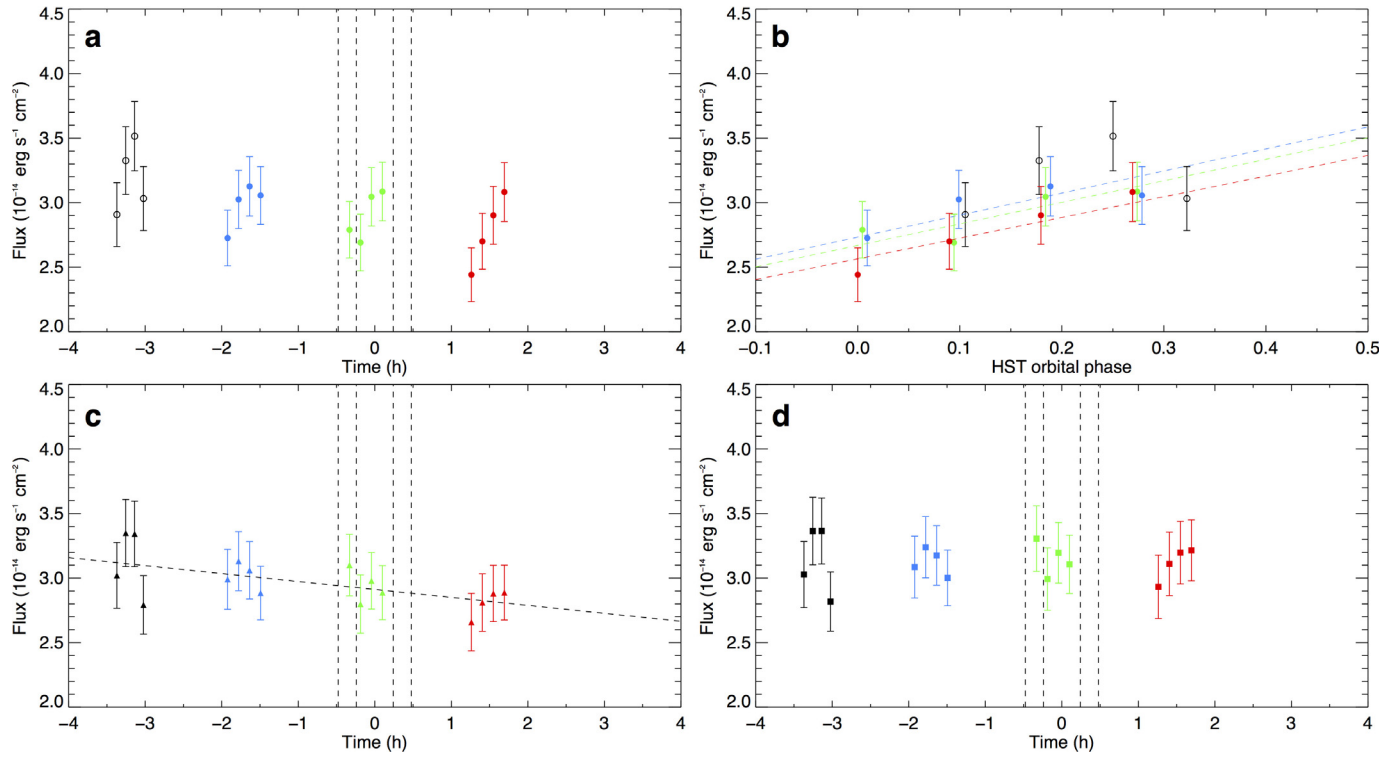
**Extended Data Figure 3 | Light curve from visit 1 data.** **a, b,** The light curve was calculated from integration of the flux over the red part of the line in the velocity range  $[+20, +250]$  km s $^{-1}$ . The temporal (**a**) and phase-folded (**b**) light curves are shown. The different colours represent data acquired during different consecutive HST orbits. The vertical dashed lines in **a** indicate the location of the optical transit contacts. The horizontal dashed lines in **b** show the best-fit constants to each HST orbit (orbit 1 data—the open circles—are not

fitted because of its different phasing). We did not apply any correction to this visit. We also did not find it necessary to trim out the first HST orbit in this visit (or in the subsequent ones) due to possible increased systematics, as described previously<sup>18</sup>. These time series have a higher sampling (by a factor of  $\sim 2$ ) than those in Fig. 2, which is made possible by exploiting the time-tag mode of data acquisition. The  $1\sigma$  error bars have been propagated accordingly.



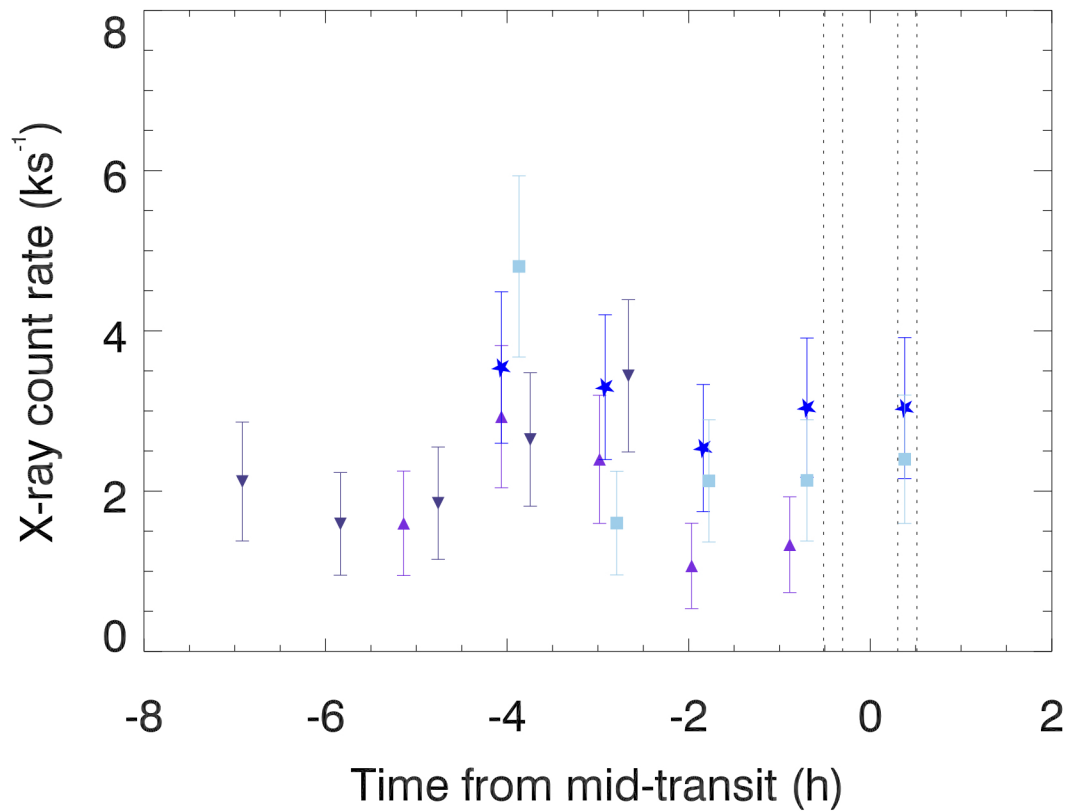
**Extended Data Figure 4 | Correction steps for the visit 2 data.** **a–d,** The different colours represent different consecutive HST orbits. The vertical dashed lines in the temporal light curves indicate the optical transit contact points. **a,** Raw light curve obtained from integration of the flux over the red part of the line in the velocity range  $[+20, +250]$   $\text{km s}^{-1}$ . **b,** Breathing correction. This is the same light curve as in **a**, phase-folded on the HST orbit. The dashed

colour lines are linear fits to the different HST orbit data. Data from HST orbit 1 (empty black circles) are not taken into account because of their slightly different phasing. **c,** Light curve corrected from telescope breathing. A linear trend (dashed line) representing long-term systematics is fitted to the data. **d,** Light curve corrected for telescope breathing and long-term systematics. All error bars are  $1\sigma$ .

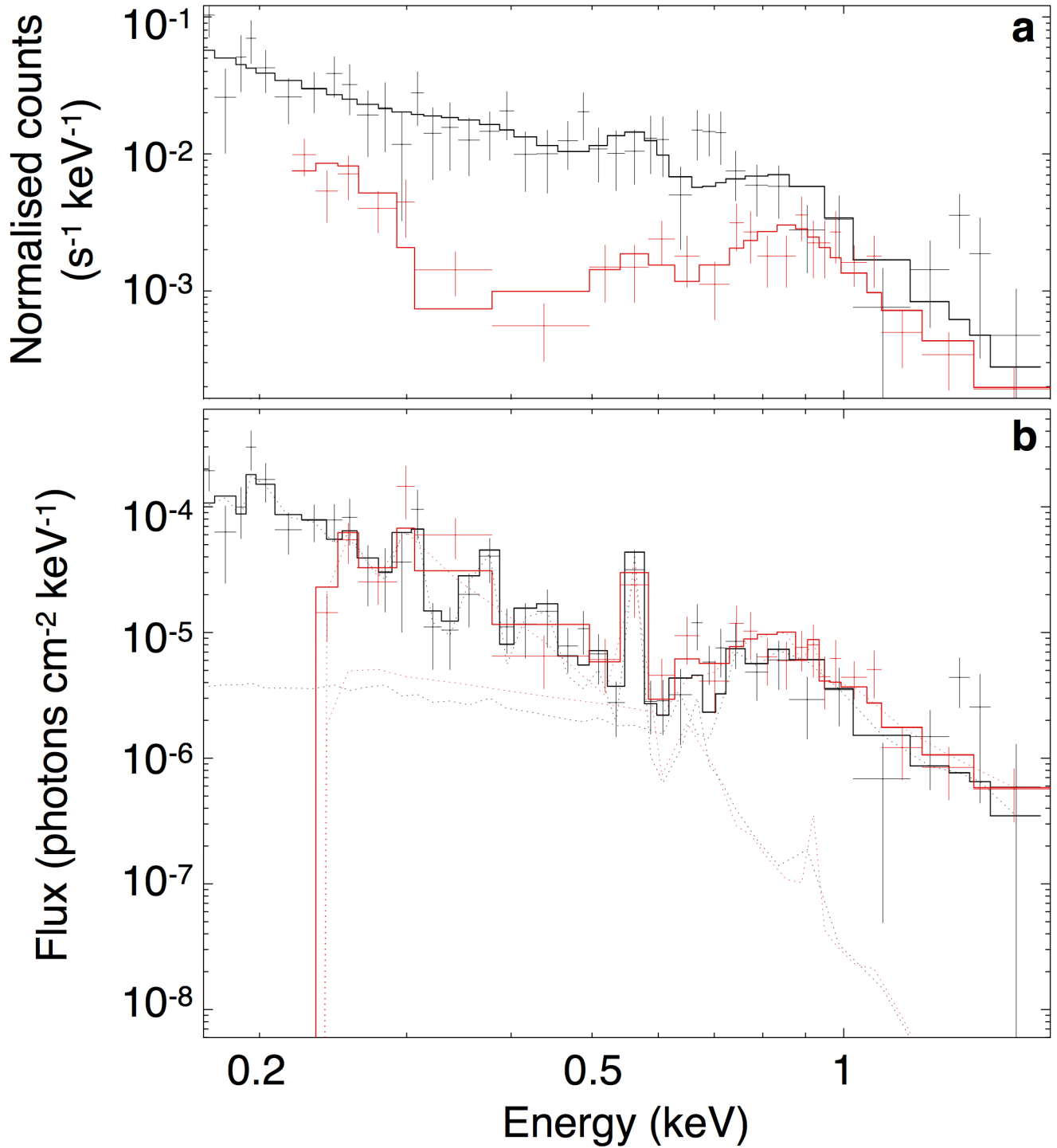


**Extended Data Figure 5 | Correction steps for the visit 3 data.** Identical to the description of Extended Data Fig. 2 for visit 2 data. All error bars are  $1\sigma$ .



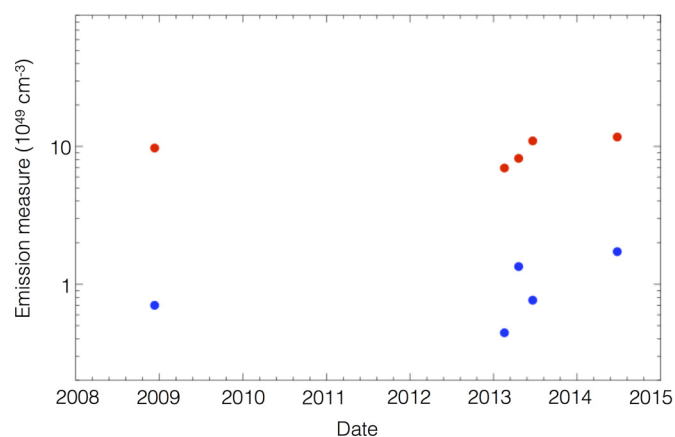


**Extended Data Figure 6 | Chandra X-ray counts of GJ 436.** The different symbols show the different visits. The June 2013 (stars) and June 2014 (squares) observations were contemporaneous with HST visits 2 and 3. The vertical dashed lines represent the contacts of the optical transit. All error bars are  $1\sigma$ .



**Extended Data Figure 7 | X-ray spectrum of GJ 436 fitted with a two-temperature model.** **a, b,** The XMM-Newton EPIC-pn spectrum is shown in black, the combined Chandra ACIS-S spectrum is shown in red. **a,** Spectra in units of measured counts, with the Chandra spectrum below XMM-Newton

because of the lower sensitivity of the instrument. **b,** Unfolded spectra in flux units; it can be seen that the XMM-Newton and Chandra spectra are consistent with each other. The dotted lines show the contributions of the individual temperature components. All error bars are  $1\sigma$ .



**Extended Data Figure 8 | Emission measures for the two temperature components fitted to the individual XMM-Newton and Chandra spectra.**

The red points correspond to the low temperature component (0.097 keV) that dominates the spectra below 0.5 keV. The blue points correspond to the higher temperature component (0.72 keV) that dominates above 0.5 keV. The higher temperature component is more variable, probably due to the effect of stellar flares. The less variable lower temperature component is a proxy for the unobserved extreme-ultraviolet emission that dominates the high-energy irradiation of the planet. The last two points are contemporaneous with HST visits 2 and 3.



Extended Data Table 1 | Log of the HST and Chandra observations

| Observation date | HST/STIS/G140M   | Chandra/ACIS-S                       |
|------------------|--|--------------------------------------|
| 2010-01-05       | GO#11817<br>1 HST orbit ("visit 0")<br>Ref. 17             |                                      |
| 2012-12-07       | GTO#12034<br>1 transit (4 HST orbits, "visit 1")<br>Ref. 6 |                                      |
| 2013-02-16       | GO#12965<br>Delayed, no data                               | Obs ID 14459<br>1 transit (18.77 ks) |
| 2013-04-18       | GO#12965<br>Target pointing problem, no data               | Obs ID 15537<br>1 transit (19.75 ks) |
| 2013-06-18       | GO#12965<br>1 transit (4 HST orbits, "visit 2")            | Obs ID 15536<br>1 transit (18.91 ks) |
| 2014-06-23       | GO#12965<br>1 transit (4 HST orbits, "visit 3")            | Obs ID 15642<br>1 transit (18.77 ks) |

Extended Data Table 2 | Ephemerides of GJ 436b transit

| Mid-transit time $T_0$<br>(JD) | Period $P$<br>(days) | Reference |
|--------------------------------|----------------------|-----------|
| BJD 2 456 295.431 924(45)      | 2.643 897 82(8)      | Ref. 31   |
| BJD 2 454 865.084 034(35)      | 2.643 898 03(30)     | Ref. 32   |
| BJD 2 454 865.083 208(42)      | 2.643 897 9(3)       | Ref. 26   |
| HJD 2 454 415.620 74(8)        | 2.643 890 4          | Ref. 30   |
| HJD 2 454 415.620 74(8)        | 2.643 850(90)        | Ref. 6    |

The offset between BJD and heliocentric Julian days (HJD) is small, typically  $\pm 4$  s.

\* Values used in the present work.

# Linking high harmonics from gases and solids

G. Vampa<sup>1</sup>, T. J. Hammond<sup>1</sup>, N. Thiré<sup>2</sup>, B. E. Schmidt<sup>2</sup>, F. Légaré<sup>2</sup>, C. R. McDonald<sup>1</sup>, T. Brabec<sup>1</sup> & P. B. Corkum<sup>1,3</sup>

**When intense light interacts with an atomic gas, recollision between an ionizing electron and its parent ion<sup>1</sup> creates high-order harmonics of the fundamental laser frequency<sup>2</sup>. This sub-cycle effect generates coherent soft X-rays<sup>3</sup> and attosecond pulses<sup>4</sup>, and provides a means to image molecular orbitals<sup>5</sup>. Recently, high harmonics have been generated from bulk crystals<sup>6,7</sup>, but what mechanism<sup>8–12</sup> dominates the emission remains uncertain. To resolve this issue, we adapt measurement methods from gas-phase research<sup>13,14</sup> to solid zinc oxide driven by mid-infrared laser fields of 0.25 volts per ångström. We find that when we alter the generation process with a second-harmonic beam, the modified harmonic spectrum bears the signature of a generalized recollision between an electron and its associated hole<sup>11</sup>. In addition, we find that solid-state high harmonics are perturbed by fields so weak that they are present in conventional electronic circuits, thus opening a route to integrate electronics with attosecond and high-harmonic technology. Future experiments will permit the band structure of a solid<sup>15</sup> to be tomographically reconstructed.**

Fifty years ago, Keldysh introduced an important idea. Treating both gases and solids on the same footing, he showed that strong-field ionization—valence-band to conduction-band transitions in solids—can be approximated by tunnelling<sup>16</sup>. Over the following decades, research on the interaction of intense pulses with solids and gases has diverged. Much of the solid-state research has been motivated by laser materials processing<sup>17</sup>. A great deal of the gas-phase research has been motivated by high-harmonic generation<sup>18,19</sup>, a process initiated by tunnelling and completed by the recollision of a highly energetic electron with its associated ion (hole). However, the deep similarity in their response to strong fields has been largely ignored. Now that high harmonics have been measured from ZnO (ref. 6) and other condensed media<sup>7,20</sup>, it is important to reconsider the relation between the gas-phase and solid-state mechanisms.

Individual ionizing atoms provide two sources of harmonics. Expressed in the language of solid-state studies, there is a ‘single-band’ (intraband) contribution, in which the recently tunnelled electron wave packet undergoes oscillatory motion in the laser field. Here the nonlinearity enters through the step-wise nature of tunnel ionization<sup>21</sup>. In addition, there is a ‘two-band’ (interband) source. In atoms, this source is interpreted in terms of the ionized electron recolliding and recombining with its associated hole (parent ion), emitting a high-energy photon in the process. In gases, the former source dominates the low harmonics while the latter source dominates high-harmonic generation. Electrons and holes in solids exhibit similar mechanisms as they move in, and between, the conduction and valence bands<sup>10,11</sup>. However, in a solid, the electron’s interaction with the lattice makes its oscillation nonlinear; this is a new source of intraband emission<sup>8</sup>, and leads to a generalized recollision<sup>11</sup>. The relative contribution of these intra- and interband mechanisms and their interplay is important theoretically<sup>8–12</sup> and experimentally<sup>7</sup>. The unresolved issue is under what circumstances does one physical mechanism dominate production of high harmonics in solids.

To resolve this issue, we adapt a method of gas-phase research to solids<sup>13</sup>. Using a mid-infrared fundamental laser to generate odd

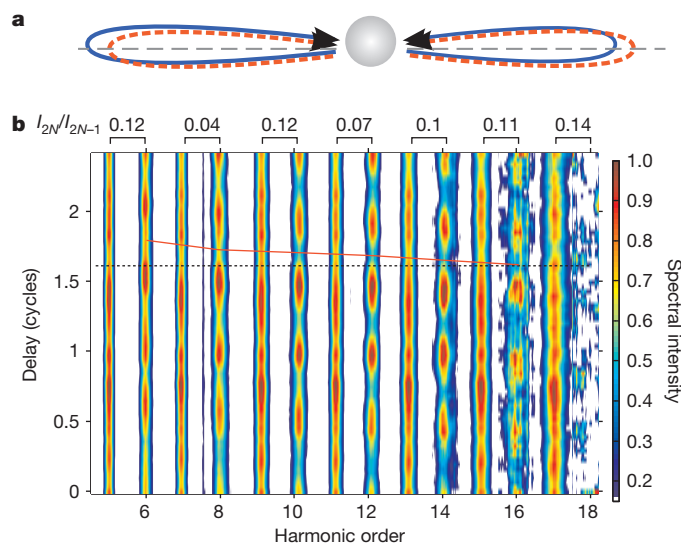
harmonics in ZnO, we perturb the process with a very weak second-harmonic beam, thereby producing even harmonics. The strength of the even harmonics modulates as the second harmonic is delayed relative to the fundamental. The phase of the modulation for each harmonic order determines the spectral phase of the emitted harmonic beam and characterizes the generation mechanism. Using photons with energies well below that of resonance ( $h\nu \approx 0.1E_g$ , where  $E_g$  is the bandgap of the solid), we find that a generalized recollision between the electron and its associated hole plays the dominant role in high-harmonic emission from ZnO at the field strength of our experiment, as it does in gases. By ‘generalized recollision’ we mean that between tunnelling and recollision the electron and hole move on their respective bands, as discussed in the Methods section. Furthermore, as we increase the relative second-harmonic intensity, for each intensity range we find solid-state behaviour that closely mirrors the atomic response. There is, however, one possible exception. As described in the Methods section, in gases there are so-called long and short trajectory contributions, which lead to differently diverging radiation. Although we have searched for them, we have not resolved long-trajectory harmonics in our experiment. However, even this may be the same in gases since long-trajectory harmonics are hard to observe with infrared drivers.

To understand the experiment it is useful to think of high-harmonic generation as a balanced interferometer. As illustrated in Fig. 1a for atomic recollision, the electron (and hole) wave packets created on subsequent half-cycles of the driving field propagate on equal but oppositely directed trajectories (equal-length arms of an interferometer), creating sub-cycle bursts of short-wavelength radiation when they recollide. With each subsequent pulse having opposite phase because of the oppositely directed collision, when the pulses are spectrally resolved, only odd harmonics constructively interfere. Adding a weak second-harmonic field unbalances the trajectories (interferometer), adding a small phase to one arm while removing it from the other, thereby producing even harmonics. For ZnO, because of the low reduced carrier mass near the band centre, low-photon-energy harmonics are very sensitive to weak second-harmonic fields. In a recollision process, the phase that maximally unbalances the interferometer differs between harmonic orders. While illustrated for recollision, the interferometer analogy also applies to intraband oscillations: here, the asymmetry originates from the slightly different current generated in two successive half-cycles, as a result of the electron reaching higher (lower) momentum when it propagates in the left (right) side of the Brillouin zone. Because harmonics are emitted as the electron moves, they are all phased together<sup>10</sup>. The second harmonic perturbs all of them equally and simultaneously, so the phase that maximally unbalances the interferometer is the same for all harmonics. This different behaviour allows us to discriminate between recollision and non-recollision based mechanisms.

Experimentally, high harmonics are produced by focusing the fundamental and the second harmonic of mid-infrared laser pulses (3.76  $\mu\text{m}$  central wavelength) in a 500-nm ZnO single crystal at an intensity of  $0.85 \text{ TW cm}^{-2}$  ( $0.25 \text{ V Å}^{-1}$ ) in the crystal. The two colours

<sup>1</sup>Department of Physics, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada. <sup>2</sup>INRS-EMT, 1650 boulevard Lionel-Boulet, CP 1020, Varennes, Québec J3X 1S2, Canada. <sup>3</sup>National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada.





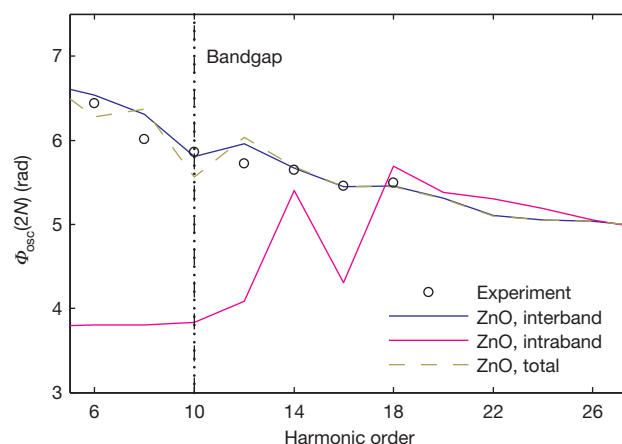
**Figure 1 | Realization of the *in situ* experiment.** **a**, Pictorial representation of the electron's trajectories. In the fundamental field alone, the electron travels symmetrically to the left and to the right of the ion (central grey sphere) during two successive half-cycles (red dashed lines). The addition of the second harmonic lengthens one trajectory while shortening the other (solid blue lines), thereby breaking the symmetry and producing even harmonics. **b**, Measured high-harmonic spectra versus the delay between the two colours; the spectral intensity is colour-coded (key at right). The even harmonic intensity modulates every half-cycle. The odd harmonics also modulate, but the modulation is much weaker. The phase of the even harmonic modulation depends on the harmonic order, as evidenced by the red solid line relative to the black reference line. (The red line links the minima in the modulation of the even harmonic intensity.) Each harmonic order is normalized separately. The relative intensity between even-order and adjacent odd-order harmonics is given at the top ( $I_{2N}/I_{2N-1}$ ). The delay of the second harmonic is defined in cycles of the fundamental.

(that is, fundamental and second harmonic) have parallel polarizations (details of the experimental set-up are given in the Methods section).

Figure 1b shows the harmonic spectrum versus the relative phase between the two colours, for a second-harmonic intensity  $9 \times 10^{-6}$  times that of the fundamental. Each harmonic order is independently normalized; the relative intensity between adjacent orders ( $I_{2N}/I_{2N-1}$ ) is reported at the top of the graph ( $N$  is an integer that identifies the harmonic order). The weak even-harmonic signal modulates with an order-dependent phase  $\Phi_{\text{osc}}(2N)$  as a function of delay. Similar behaviour is measured for atomic harmonics<sup>13</sup>, where  $\Phi_{\text{osc}}$  is used to extract the spectral phase of each even harmonic or, equivalently, the emission time of the attosecond pulses. Here, we use  $\Phi_{\text{osc}}$  to establish the origin of solid-state harmonics.

Figure 2 shows  $\Phi_{\text{osc}}$  plotted as a function of harmonic order (black circles) up to the cut-off of our spectrograph. We use an unmeasured constant phase to position the experimental data (this constant phase is measurable using nonlinear optical methods). The theoretical values of  $\Phi_{\text{osc}}$  calculated from a two-band model are also plotted for the intraband (purple line) and interband (blue line) sources, and for their combined emission (gold dashed line); the calculations are described in the Methods section. We find that the experiment is consistent with the theoretical prediction (Extended Data Fig. 6) that interband harmonics dominate in ZnO for the field strength used (we note that recent work has found evidence for the dominance of intraband oscillations for THz excitation<sup>7</sup>). In the Methods section we investigate theoretically how our results scale with the field strength of the fundamental.

The slope of the curves in Fig. 2 is determined by both the intensity of the fundamental and the band structure of the crystal. Crystal dispersion has two major consequences: first, the electron-hole pair accumulates more phase than in a gas ( $\sim 10$  times for the short trajectories), making it easier to break the symmetry with the second-harmonic



**Figure 2 | Comparison with theory.** The phase of the modulation of the even harmonics is extracted from Fig. 1 (black circles) and compared to the simulated intraband (purple line) and interband (blue line) phase. The simulated phase for their combined emission (intraband plus interband, yellow dashed line) agrees with the interband emission, in agreement with interband harmonics being significantly stronger than intraband ones (Extended Data Fig. 6).

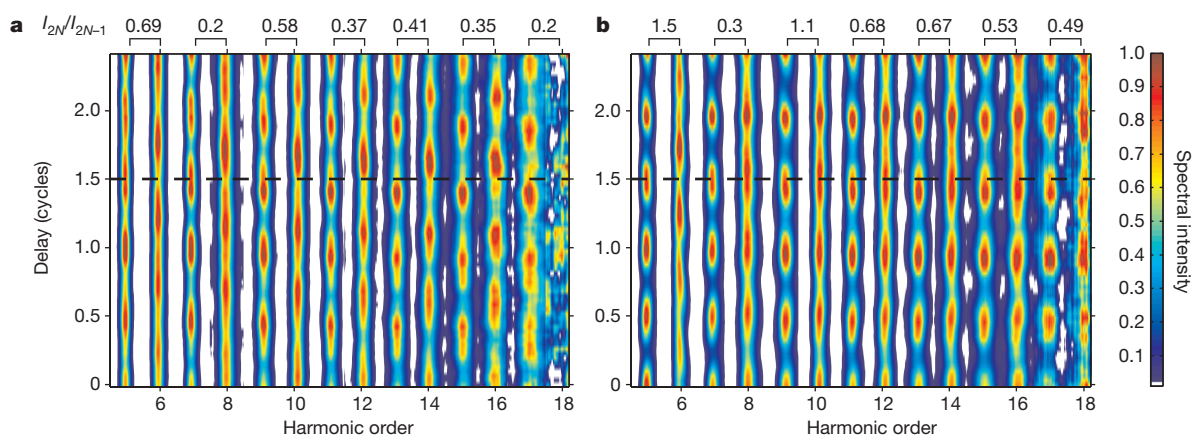
field. This is why we require second-harmonic intensities of only  $\sim 10^{-5}$  times that of the fundamental. Second, crystal dispersion influences the duration of the high-harmonic pulses. We measure a harmonic chirp of  $0.38 \text{ fs eV}^{-1}$  at the 16th harmonic, corresponding to a train of pulses with duration of 1.7 fs (before dispersion compensation) for a  $\sim 2 \text{ eV}$  bandwidth. Wider bands and higher intensities will result in smaller chirps, thereby allowing attosecond pulse generation from solids.

As the second-harmonic intensity is increased, we move from perturbing to controlling harmonic emission. In gases, we first reach an intensity where  $\Phi_{\text{osc}}$  is nearly out of phase for the even and odd harmonics<sup>13</sup>. Figure 3a shows experimental results for ZnO using a second-harmonic intensity of  $\sim 10^{-4}$  times that of the fundamental. The alternation between even and odd harmonics occurs when the second harmonic modulates the electron phase by  $\Phi_{\text{osc}} = \pi/2$ . For even higher second-harmonic field strengths (Fig. 3b), high-harmonic emission is substantially modified. In the highly asymmetric sum field, tunnelling only occurs once per laser cycle. This leads to simultaneous emission (or suppression) of even and odd harmonics—a behaviour well-studied for atomic gases<sup>14</sup>.

Taken together, our findings provide strong evidence that, for the field strength and laser wavelength of our experiment, a generalized recollision between an electron and its associated hole is the primary source of high harmonics in ZnO. Thus, 50 years after Keldysh's seminal paper<sup>16</sup>, we rebuild the connection linking the high-field response of gases and solids. But decoherence is a major difference between solids and gases. The occurrence of electron-hole recombination demonstrates the tolerance of the short-lived electron to dephasing. Although scattering times vary considerably between materials<sup>22,23</sup> and with excitation conditions<sup>24</sup>, our measurements suggest that high-harmonic technology could have a broad impact in condensed media.

In this paper we have exploited our ability to manipulate the phase that the electron-hole pair acquires to identify the fundamental mechanism dominating harmonic generation in ZnO. The phase also encodes information about the band structure of the material in which the electron travelled. By measuring  $\Phi_{\text{osc}}$  as a function of crystal orientation, it will be possible to reconstruct a material's three-dimensional momentum-dependent band structure<sup>15</sup>. This all-optical technique will allow measurement of band structures where it is not possible to detect photoelectrons, for example in high-pressure experiments<sup>25</sup>.

Although we have not concentrated on it here, the tunnelling step determines the harmonic intensity. Our weak second-harmonic beam



**Figure 3 | From perturbing to controlling harmonic generation.** **a**, When the intensity of the second harmonic reaches  $1 \times 10^{-4}$  times that of the fundamental, the even and odd harmonics modulate out of phase. **b**, Further increasing the intensity to  $3 \times 10^{-3}$  times that of the fundamental, we find that

not only modifies the electron phase but also slightly alters tunnelling, thereby leading to oscillations of the odd harmonics with the delay between the two colours. The relative phase of the modulation between even and odd orders contains detailed information on the time-dependent tunnelling rate<sup>26</sup>. Applying this analysis to solids will allow us to probe the sub-cycle dynamics of strong-field tunnelling in solids<sup>27,28</sup>—the essence of laser material modification<sup>17</sup>.

Finally, independent of the mechanism, we have demonstrated that a clear spectral signature can be imposed on a high-harmonic beam using a control beam having a peak electric field of  $\sim 5 \text{ V } \mu\text{m}^{-1}$ . Fields of this magnitude are present in electronic circuits<sup>29</sup>. Thus, we have bridged the gap between electronics and attosecond physics. If the sensitivity that we have demonstrated to perturbing fields can be transferred to silicon, it will be possible to record movies of working semiconductor electrical circuits or plasmons propagating in nano-plasmonic devices. Harmonics generated from a spatially complex circuit will diffract according to the instantaneous distribution of the internal fields. The control field can also be applied by electrodes on the crystal, thereby allowing manipulation of the harmonic beam with DC or pulsed electrical signals. Thus we have taken a significant step towards a new area of research, that of solid-state attosecond electronics<sup>30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 November 2014; accepted 27 April 2015.

- Corkum, P. B. A plasma perspective on strong field multiphoton ionization. *Phys. Rev. Lett.* **71**, 1994–1997 (1993).
- L'Huillier, A. & Balcou, P. High-order harmonic generation in rare gases with a 1-ps 1053-nm laser. *Phys. Rev. Lett.* **70**, 774–777 (1993).
- Spielmann, Ch. *et al.* Generation of coherent soft X-rays in the water window using 5-femtosecond laser pulses. *Science* **278**, 661–664 (1997).
- Goulielmakis, E. *et al.* Single-cycle nonlinear optics. *Science* **320**, 1614–1617 (2008).
- Itatani, J. *et al.* Tomographic imaging of molecular orbitals. *Nature* **432**, 867–871 (2004).
- Ghimire, S. *et al.* Observation of high-order harmonic generation in a bulk crystal. *Nature Phys.* **7**, 138–141 (2011).
- Schubert, O. *et al.* Sub-cycle control of terahertz high-harmonic generation by dynamical Bloch oscillations. *Nature Photon.* **8**, 119–123 (2014).
- Kemper, F., Moritz, B., Freericks, J. K. & Devereaux, T. P. Theoretical description of high-order harmonic generation in solids. *New J. Phys.* **15**, 023003 (2013).
- Higuchi, T., Stockman, M. I. & Hommelhoff, P. Strong-field perspective on high-harmonic radiation from bulk solids. *Phys. Rev. Lett.* **113**, 213901 (2014).
- Golde, D., Meier, T. & Koch, S. High harmonics generated in semiconductor nanostructures by the coupled dynamics of optical inter- and intraband excitations. *Phys. Rev. B* **77**, 075330 (2008).

all harmonics modulate in phase. Each harmonic order is normalized separately. The relative intensity between even-order and adjacent odd-order harmonics is given at the top.

- Vampa, G. *et al.* Theoretical analysis of high harmonic generation in solids. *Phys. Rev. Lett.* **113**, 073901 (2014).
- Hawkins, P. G., Ivanov, M. Y. & Yakovlev, V. S. Effect of multiple conduction bands on high-harmonic emission from dielectrics. *Phys. Rev. A* **91**, 013405 (2015).
- Dudovich, N. *et al.* Measuring and controlling the birth of attosecond XUV pulses. *Nature Phys.* **2**, 781–786 (2006).
- He, X. *et al.* Interference effects in two-color high-order harmonic generation. *Phys. Rev. A* **82**, 033410 (2010).
- Ulstrup, S. *et al.* Ultrafast dynamics of massive Dirac fermions in bilayer graphene. *Phys. Rev. Lett.* **112**, 257401 (2014).
- Keldysh, L. Ionization in the field of a strong electromagnetic wave. *Sov. Phys. JETP* **20**, 1307–1314 (1965).
- Stuart, B. C. *et al.* Nanosecond-to-femtosecond laser-induced breakdown in dielectrics. *Phys. Rev. B* **53**, 1749–1761 (1996).
- Wörner, H. J., Bertrand, J. B., Kartashov, D. V., Corkum, P. B. & Villeneuve, D. M. Following a chemical reaction using high-harmonic interferometry. *Nature* **466**, 604–607 (2010).
- Smirnova, O. *et al.* High harmonics interferometry of multi-electron dynamics in molecules. *Nature* **460**, 972–977 (2009).
- Zaks, B., Liu, R. B. & Sherwin, M. S. Experimental observation of electron-hole recollisions. *Nature* **483**, 580–583 (2012).
- Brunel, F. Harmonic generation due to plasma effects in a gas undergoing multiphoton ionization in the high-intensity limit. *J. Opt. Soc. Am. B* **7**, 521–526 (1990).
- Schultze, M. *et al.* Attosecond band-gap dynamics of silicon. *Science* **346**, 1348–1352 (2014).
- Schultze, M. *et al.* Controlling dielectric with the electric field of light. *Nature* **493**, 75–78 (2012).
- Vu, Q. T. *et al.* Light-induced gaps in semiconductor band-to-band transitions. *Phys. Rev. Lett.* **92**, 217403 (2004).
- Eremets, M. I. & Troyan, I. A. Conductive dense hydrogen. *Nature Mater.* **10**, 927–931 (2011).
- Shafir, D. *et al.* Resolving the time when an electron exits a tunnelling barrier. *Nature* **485**, 343–346 (2012).
- Gertsvolf, M., Spanner, M., Rayner, D. M. & Corkum, P. B. Demonstration of attosecond ionization dynamics inside transparent solids. *J. Phys. At. Mol. Opt. Phys.* **43**, 131002 (2010).
- Schiffrin, A. *et al.* Optical-field induced currents in dielectrics. *Nature* **493**, 70–74 (2013).
- Schwierz, F. Graphene transistors. *Nature Nanotechnol.* **5**, 487–496 (2010).
- Krausz, F. & Stockman, M. I. Attosecond metrology: from electron capture to future signal processing. *Nature Photon.* **8**, 205–213 (2014).

**Acknowledgements** We thank A. Laramée from the Advanced Laser Light Source for technical support during the experiment and M. Clerici for lending some equipment. We acknowledge financial support from the US AFOSR, NSERC, FRQNT, MDEIE and CFI.

**Author Contributions** G.V. and P.B.C. conceived the experiment; G.V., T.J.H. and N.T. performed the experiment; N.T. and B.E.S. developed the laser source; P.B.C. and F.L. supervised the experiment; C.R.M. and T.B. supervised the theoretical calculations; all authors contributed to the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.B.C. ([paul.corkum@nrc.ca](mailto:paul.corkum@nrc.ca)) or G.V. ([gvamp015@uottawa.ca](mailto:gvamp015@uottawa.ca)).

## METHODS

**Experimental set-up.** High harmonics are generated with a mid-infrared laser source centred around 3.76  $\mu\text{m}$  and with 19  $\mu\text{J}$  pulse energy. The pulse duration of 95 fs is measured with a dispersion-free SHG FROG after the ZnO crystal. The mid-infrared wavelength is obtained by difference-frequency generation in a 400  $\mu\text{m}$  thick AgGaS<sub>2</sub> (AGS) crystal between the signal and the idler of a commercial Optical Parametric Amplifier (Light Conversion Topas-HE) pumped by 4.5 mJ from a Ti-sapphire laser. The output of the OPA is further amplified in a beta barium borate (BBO) crystal pumped by 10 mJ of 800 nm pulses. The mid-infrared beam is spatially filtered with a pin-hole and focused with an  $f/30$  spherical Ag mirror onto an epitaxially grown 500 nm thin film of a single crystal of wurtzite ZnO(0001) deposited on a 0.5 mm sapphire(0001) substrate. The optical axes of both crystals are aligned parallel to the laser  $k$ -vector. We used the knife-edge technique to measure the size of the focal spot.

The second harmonic is generated in a 300  $\mu\text{m}$  AGS crystal optimized for type-I SHG right after the pin-hole to exploit the high intensity and high beam quality. The second-harmonic intensity is controlled by moving the AGS crystal closer to or farther from the pin-hole. The two beams are separated and then recombined before the focusing mirror with dichroic beam splitters (LaserOptick) that reflect the second harmonic. Its polarization is rotated with a broadband  $\lambda/2$  plate before recombination. The delay between the two colours is scanned with a PZT stage on the second-harmonic arm.

The high harmonics are refocused with an Al mirror into a visible-ultraviolet spectrometer from OceanOptics.

**The theoretical model.** To interpret the experiment, we use the theoretical framework developed in ref. 11. The nonlinear response of the solid is described by a two band model whose characteristics are chosen to be those of ZnO. The model solves the semiconductor Bloch equations for the time-dependent band populations  $n_m(\mathbf{k}, t)$  ( $m = v, c$  for the valence and conduction bands respectively) and for the interband polarization  $p(\mathbf{k}, t)$ , where  $\mathbf{k}$  is the crystal momentum, for a single electron-hole pair. Coulomb interaction between different pairs, leading to coupling of different  $k$ -states<sup>31</sup>, is neglected. The band parameters and lattice constants are reported in ref. 32. The laser parameters are: frequency  $\omega = 0.0121$  a.u., peak field strength  $F_0 = 0.0049$  a.u.

Extended Data Fig. 1 shows the population of the conduction band along the  $\Gamma\text{M}$  direction of the reciprocal space as a function of time. The laser field creates electron-hole pairs around each field extreme. They are then accelerated to high crystal momenta and back to the centre of the Brillouin zone in a process that repeats every half laser cycle.

The model naturally includes the sub-cycle exchange of the band population—the analogue of the step-wise ionization in gases. However, in solids both the single-band (intraband) and the two-band (interband) mechanism bear the signature of the generalized electron-hole recollision.

Extended Data Fig. 2a, b shows the high-harmonic spectrum of the respective intraband and interband terms as a function of time. The spectra are obtained from a windowed Fourier transform of the intra- and interband currents, defined in ref. 11, which can be separately calculated by the model. A 0.35 cycles wide Blackman window is scanned across half a laser cycle. The resulting spectra are continuous because only one recollision event is allowed inside the temporal window. We apply a super-Gaussian spectral filter that does not affect above bandgap harmonics, but progressively filters lower harmonics. This is done because the temporal window is too narrow for the lower orders (and the fundamental frequency) and results in artificial broadening of these harmonics which ultimately masks the weak high-harmonic signal if the filter is not applied. Interband harmonics above the bandgap are approximately  $10^4$  times stronger than intraband ones. For each harmonic photon energy, the maximum spectral intensity determines its time-of-emission (or spectral phase).

For comparison, the time of generalized recollision is traced in red in Extended Data Fig. 2a, b. The time of re-encounter of the electron with its associated hole is extracted from their classical motion in the conduction (for the electron) and valence (for the hole) bands relative to the time of the field crest in which they were created at zero crystal momentum by strong-field tunnelling. Their acceleration in reciprocal space is  $\mathbf{k}(t) = \mathbf{A}(t) - \mathbf{A}(t')$ , where  $\mathbf{A}(t)$  is the laser vector potential, and  $t'$  is the time of creation of the electron-hole pair. The real space velocity is  $\mathbf{v}_m[\mathbf{k}] = \nabla_{\mathbf{k}} E_m(\mathbf{k})$ , with  $E_m(\mathbf{k})$  the energy dispersion for band  $m$ , from which the instantaneous position is obtained by temporal integration. This semi-classical model, introduced in ref. 11, has been detailed in ref. 32.

Extended Data Fig. 2 shows that both interband and intraband currents produce high harmonics with a spectral phase determined by the generalized recollision. This recollision contribution is contained in the structure of the intraband equation<sup>11</sup>. Contrary to interband harmonics, intraband emission is dominated by nonlinear oscillations and by tunnelling below the 18th harmonic order. The transition is reflected in the *in situ* calculation (see Fig. 2), where  $\Phi_{\text{osc}}$  becomes

independent on the delay between fundamental and second harmonic below the 18th harmonic.

Many electron-hole pairs are created by the strong laser field. Owing to the interaction of one electron (hole) with the others, the electron (hole) can scatter to a different  $k$ -state than that of its correlated hole (electron), therefore preventing their recombination. However, in a many-body theory, each carrier can encounter a different oppositely charged partner. In atomic high harmonic generation it was argued that recombination of the electron with an uncorrelated hole is not allowed in principle<sup>33</sup>. This debate is yet to be settled in solids. Our experiment only detects recollision between correlated electron-hole pairs.

**Modelling the experiment.** A weak second-harmonic field is added to the fundamental one to simulate the *in situ* experiment. The high-harmonic spectrum is recorded for each phase between the two colours. The resulting spectrograms for the interband and intraband emission are reported in Extended Data Fig. 3a, b, respectively. We require the second harmonic to be  $6 \times 10^{-4}$  relative to the fundamental to produce even high harmonics that are  $\sim 5\%$  of the odd ones. The modulation depth is close to 100% for both intra- and interband generated above bandgap harmonics. It drops to  $\sim 50\%$  for below bandgap interband harmonics and for harmonics below the 18th for intraband emission. The spectrograms of the interband and intraband currents and of the sum of the two are analysed in the same way as the experimental data: (i) the intensity of each harmonic is spectrally integrated over its width; (ii) the resulting intensity modulation as a function of delay is offset to 0 and normalized to 1 and (iii) fitted to a cosine function of delay with form  $\cos(\phi_{\text{delay}} + \Phi_{\text{osc}})$ . Finally,  $\Phi_{\text{osc}}$  for both emission mechanisms and for the total emission (the sum of the two) is extracted and shown in Fig. 2 of the main text.

**Comparison with atomic high-harmonic generation.** We analyse the similarities between the high harmonics from atoms and solids based on calculation of classical trajectories. In Extended Data Fig. 4 we plot  $\Phi_{\text{osc}}$  calculated for an isolated atom (green line). The field strength of the fundamental is increased to  $0.44 \text{ V \AA}^{-1}$  to allow the same cut-off as in ZnO. The phase for interband emission in ZnO and the experimental data are the blue line and the black circles, respectively. According to the analytical equivalence between atomic and solid emission, demonstrated in refs 11, 32, the intensity modulation of the even harmonics can be calculated starting from the additional dipole phase<sup>13</sup>:

$$\sigma(t, \phi) = \int_{t'(t)}^t \mathbf{v}(\tau, t') \mathbf{A}_2(\tau, \phi) d\tau$$

Where  $\mathbf{A}_2(\tau, \phi) = \mathbf{A}_2 \cos(2\omega\tau + \phi)$  is the vector potential of the second harmonic and  $\mathbf{v}(\mathbf{k}) = \nabla_{\mathbf{k}} \varepsilon_g(\mathbf{k})$ , with  $\varepsilon_g(\mathbf{k})$  the momentum dependent bandgap. Once the semi-classical trajectories are known (as discussed in the previous section), the integral can be computed and the phase of the modulation extracted from<sup>13</sup>  $\Phi_{\text{osc}} = \text{atan}(\sigma_c/\sigma_s)$  with  $\sigma(t, \phi) = \sigma_s(t) \cos(\phi) + \sigma_c(t) \sin(\phi)$ .

In atomic high-harmonic generation there are two important contributions to a given high harmonic. One is from the so-called short trajectory electrons. They appear as the curve with negative slope at larger phases in Extended Data Fig. 4. Almost all research on gas-phase harmonics has concentrated on this contribution. There are also long trajectory electrons. They are responsible for the other part of the curve and are often ignored in gases since their divergence is larger and an experiment can be designed to minimize their impact.

In ZnO, the long trajectory contribution is significantly different from the atomic case. This is a direct consequence of the non-parabolic band dispersion. The distortion of the trajectories is extensively analysed in ref. 32. Further, no prediction can be made for below-bandgap harmonics, since classically the electron cannot recollide with energy smaller than the minimum bandgap. A quantum mechanical calculation that includes the effect of the tunnelling step is required to extend the comparison to smaller photon energies.

**Theoretical cut-off scaling of interband emission.** The high-harmonic spectrum extends up to a defined maximum harmonic order, called the cut-off. In atoms, the cut-off is related to the maximum kinetic energy available to the electron when it recombines with the hole. The cut-off depends linearly on the ponderomotive energy  $U_p$ , which, in turn, depends quadratically on the field strength.

In contrast with atomic harmonics, experiments in solids<sup>6</sup> show a linear scaling of the cut-off with field strength. Theoretical investigation of the intraband current confirmed the linear scaling<sup>8</sup>. Extended Data Fig. 5 shows that the simulated interband current also leads to a cut-off that scales linearly with the field strength. Exploiting the analytical solution for this current<sup>32</sup>, it is possible to derive a semi-classical approximate cut-off scaling law that fully agrees with the simulation. This analysis is reported in ref. 32.

**Role of interband emission for various field strengths.** We theoretically investigate the relative importance of interband and intraband emission at different field strengths at the laser wavelength used in this experiment. Extended Data Fig. 6 reports the interband and intraband spectra for  $F_0 = 0.003$  a.u.,  $F_0 = 0.0049$  a.u.

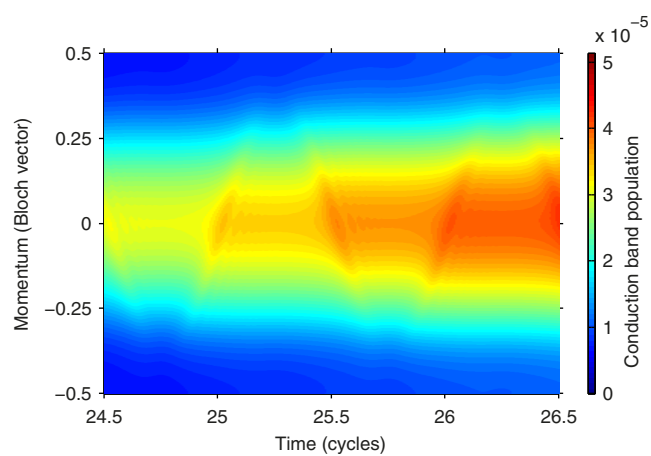
(corresponding to the field strength reached in our experiment) and  $F_0 = 0.008$  a.u. (corresponding to the field strength reached in ref. 6, including the loss from normal incidence reflection). In all cases, interband emission dominates by four orders of magnitude for harmonic orders above the minimum bandgap (marked by the vertical dashed black line). Below the minimum bandgap the difference is less.

A less clear harmonic structure is observed at the highest field strength. The loss of contrast arises from the interference between two sets of recolliding trajectories: one of electrons born slightly after the peak of the field (the 'short' branch of the atomic case) with one of electrons born just before the peak of the field. In the atomic case, electrons born before the peak do not recollide. In a solid, the

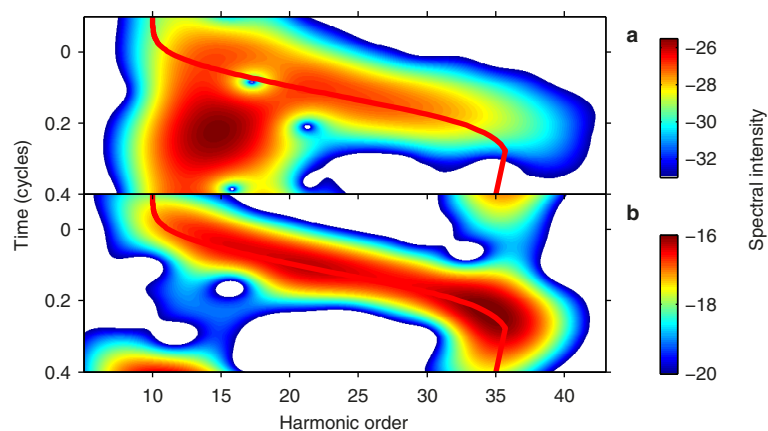
non-parabolic shape of the band structure allows these extra recombinations. The topic will be a matter of a future publication.

31. Golde, D., Kira, M., Meier, T. & Koch, S. W. Microscopic theory of the extremely nonlinear terahertz response of semiconductors. *Phys. Status Solidi B* **248**, 863–866 (2011).
32. Vampa, G., McDonald, C. R., Orlando, G., Corkum, P. B. & Brabec, T. Semiclassical analysis of high harmonic generation in bulk crystals. *Phys. Rev. B* **91**, 064302 (2015).
33. Niikura, H. *et al.* Probing molecular dynamics with attosecond resolution using correlated wave packet pairs. *Nature* **421**, 826–829 (2003).



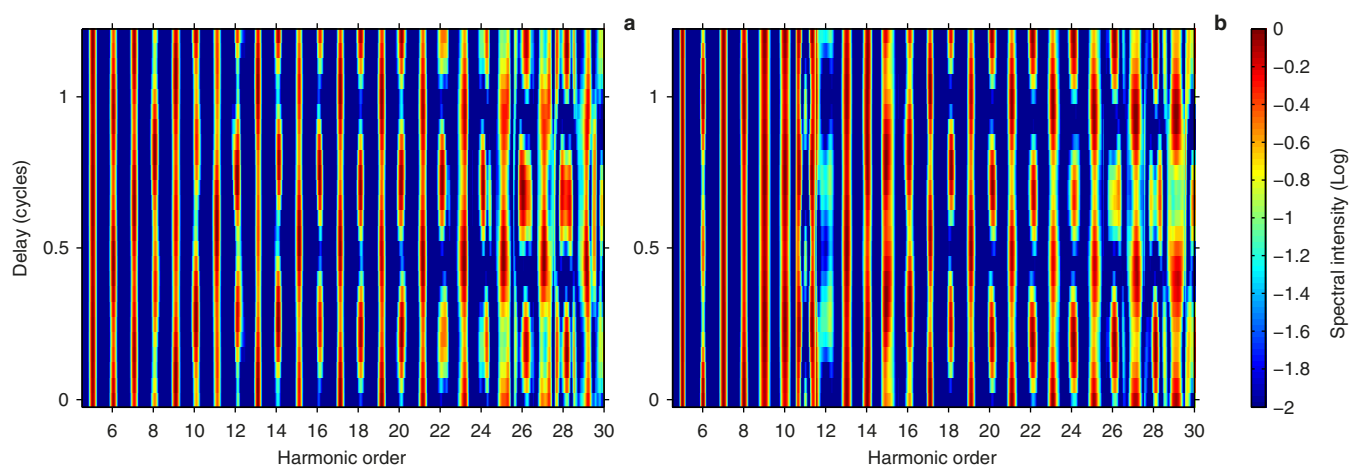


**Extended Data Figure 1 | Evolution of conduction band population with time as a function of Bloch vector.** Population in the conduction band (colour-coded, key at right) is created at each peak of the laser field and subsequently accelerated to large crystal momenta. The population is resolved along the  $\Gamma$ M direction of the Brillouin zone.

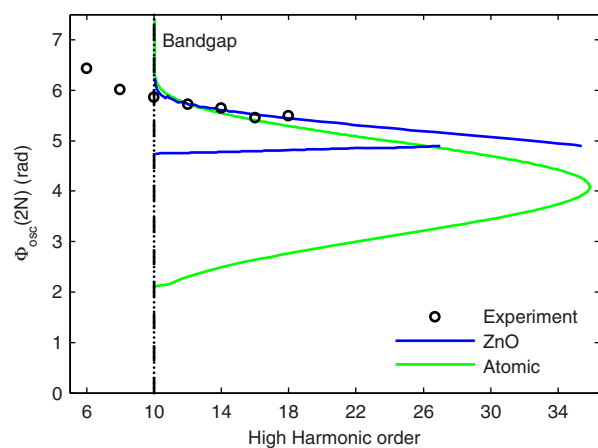


**Extended Data Figure 2 | High-harmonic emission is timed to the electron-hole recollision.** The spectral content of the intraband (a) and interband (b) mechanisms as a function of time shows that each high harmonic is emitted at a specific moment of the laser cycle. This moment is the time of generalized

recollision between the electron and the hole (red line). The intensity (colour-coded) is on a logarithmic scale. The laser parameters are:  $\omega = 0.0121$  a.u.,  $F_0 = 0.0049$  a.u.

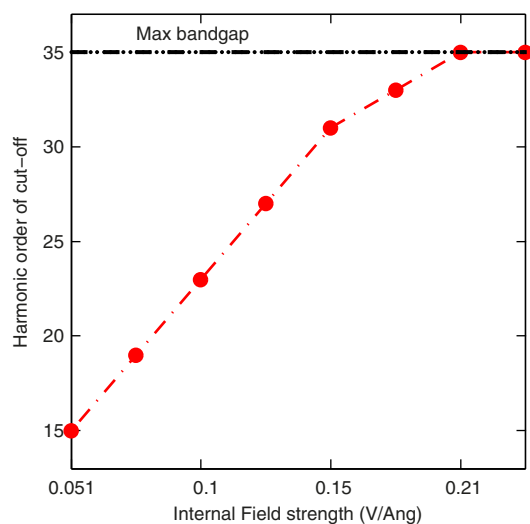


**Extended Data Figure 3 | Predicted *in situ* spectrograms.** Simulated spectrogram for interband (a) and intraband (b) emission. Intensity is colour-coded on a logarithmic scale. Each harmonic order has been normalized to simplify comparison. The laser parameters are:  $\omega = 0.0121$  a.u.,  $F_0 = 0.0049$  a.u.,  $F_{2\omega} = 10^{-2} F_0$ .

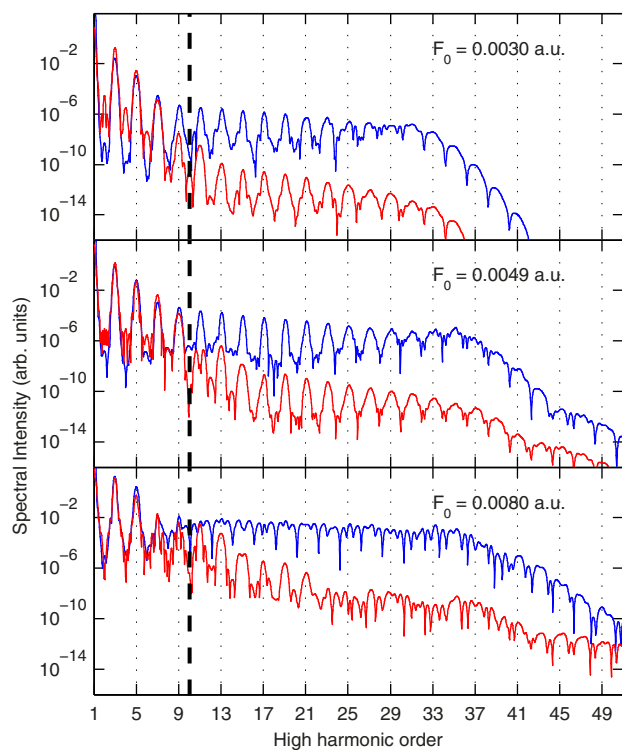


**Extended Data Figure 4 | Comparison of atomic and solid high-harmonic generation.** The phase of the modulation of the even harmonic intensity is plotted as a function of harmonic order. The green line is obtained for an atom with ionization potential equal to the bandgap of ZnO (3.3 eV, marked by the vertical black dashed line). The blue line and the black circles are the theoretical prediction for the interband source and the experimental data points, respectively. The theoretical prediction is based on classical trajectory calculation.





**Extended Data Figure 5 | Cut-off scaling of interband emission.** The high-harmonic cut-off of interband emission predicted by the numerical solution of the two-band model is plotted as a function of field strength. The cut-off scales linearly with the field strength up to the maximum bandgap (reached at the edge of the Brillouin zone). The calculated values (red dots) are connected by the dashed dotted line to show the linear scaling.



**Extended Data Figure 6 | Scaling of high-harmonic emission with the field strength of the fundamental.** The two-band model predicts that interband emission dominates over a wide range of field strengths. The field strength of our experiment is  $F_0 = 0.0049$  a.u. (middle panel), while that of ref. 6 corrected for the reflection loss is  $F_0 = 0.008$  a.u. (bottom panel). The vertical black dashed line marks the minimum bandgap (at 3.3 eV).

# Contribution of changes in atmospheric circulation patterns to extreme temperature trends

Daniel E. Horton<sup>1,2</sup>, Nathaniel C. Johnson<sup>3,4,5</sup>, Deepti Singh<sup>1</sup>, Daniel L. Swain<sup>1</sup>, Bala Rajaratnam<sup>1,2,6</sup> & Noah S. Diffenbaugh<sup>1,2</sup>

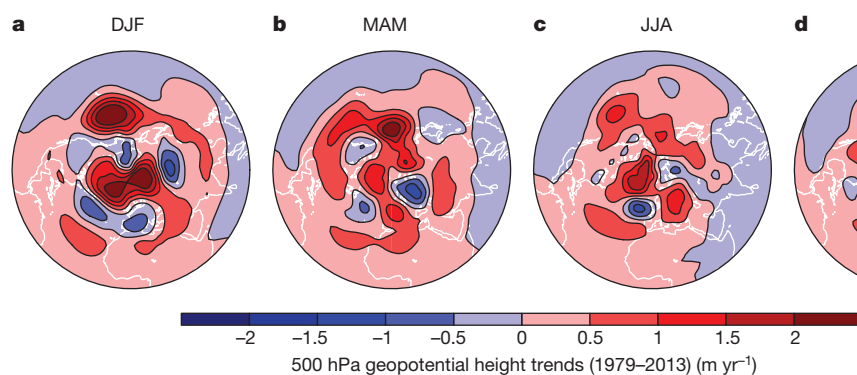
Surface weather conditions are closely governed by the large-scale circulation of the Earth's atmosphere. Recent increases in the occurrence of some extreme weather phenomena<sup>1,2</sup> have led to multiple mechanistic hypotheses linking changes in atmospheric circulation to increasing probability of extreme events<sup>3–5</sup>. However, observed evidence of long-term change in atmospheric circulation remains inconclusive<sup>6–8</sup>. Here we identify statistically significant trends in the occurrence of atmospheric circulation patterns, which partially explain observed trends in surface temperature extremes over seven mid-latitude regions of the Northern Hemisphere. Using self-organizing map cluster analysis<sup>9–12</sup>, we detect robust circulation pattern trends in a subset of these regions during both the satellite observation era (1979–2013) and the recent period of rapid Arctic sea-ice decline (1990–2013). Particularly substantial influences include the contribution of increasing trends in anticyclonic circulations to summer and autumn hot extremes over portions of Eurasia and North America, and the contribution of increasing trends in northerly flow to winter cold extremes over central Asia. Our results indicate that although a substantial portion of the observed change in extreme temperature occurrence has resulted from regional- and global-scale thermodynamic changes, the risk of extreme temperatures over some regions has also been altered by recent changes in the frequency, persistence and maximum duration of regional circulation patterns.

Although most land regions show robust warming over the past century<sup>13</sup>, the pattern of change has not been spatially uniform<sup>14</sup>. This heterogeneity results from regional differences in the response of the climate system to increasing radiative forcing, and from the background noise of climate variability. Together, these factors substantially increase the challenge of climate change detection, attribution and projection at regional and local scales<sup>14–16</sup>.

The spatial pattern of changes in extreme weather events has generated arguments that global warming has caused dynamic and/or thermodynamic changes that have differentially altered extreme event

probabilities<sup>1,17</sup>. Thermodynamic arguments are well understood and observed. For example, the accumulation of heat in the atmosphere has resulted in upward trends in hot extremes, downward trends in the majority of cold extremes, and more intense hydroclimatic events<sup>1,2</sup>. Dynamic arguments have greater uncertainties<sup>15–19</sup>. Changes in the large-scale atmospheric circulation—for instance, an increase in the occurrence or persistence of high-amplitude wave patterns—could alter the likelihood of extreme events<sup>20</sup>. Recent extremes in the Northern Hemisphere mid-latitudes<sup>1,2,17</sup> have motivated hypotheses of a dynamic linkage between ‘Arctic amplification’, altered atmospheric circulation patterns, and changes in the probability of mid-latitude extremes<sup>3–5,17</sup>. Despite divergent views on the causal direction of this linkage<sup>17</sup>, altered atmospheric dynamics are consistently invoked. Although trends in mean-seasonal mid-atmospheric geopotential heights have been identified (fig. 2.36 of ref. 21; Fig. 1), evidence of changes in the occurrence of sub-seasonal atmospheric patterns remains equivocal, as does their contribution to extreme event probabilities<sup>6–8</sup>.

Previous efforts to detect trends in atmospheric circulation may have been hampered by narrowly defined, spatially sensitive, and/or non-standardized metrics<sup>3,6–8,17</sup>. We therefore employ a large-scale spatial characterization approach—self-organizing map (SOM) cluster analysis—to track the occurrence of highly generalized mid-atmospheric circulation patterns. We use 500 hPa geopotential height anomaly fields to describe daily circulation, and group each day's pattern into one of a predefined number of SOM clusters based on a measure of pattern similarity<sup>9–12</sup> (Methods). The number of clusters is largely dependent on the degree of specificity/generalizability required to test a particular hypothesis<sup>9–12</sup>. To facilitate generalized large-scale mid-atmospheric classification, we use four clusters per domain. Using three reanalyses (Methods), we calculate linear trends ( $\text{yr}^{-1}$ ) in the time series of annual values of (1) the total number of days in each season on which each SOM pattern occurs (‘occurrence’ ( $\text{d yr}^{-1}$ )); (2) the mean length of consecutive occurrence (‘persistence’ ( $\text{d event}^{-1}$ )); and (3) the longest consecutive occurrence (‘maximum duration’ ( $\text{d event}^{-1}$ )). We consider trends in each metric to be robust if



**Figure 1 | Trends in mid-atmospheric geopotential heights.** Northern Hemisphere polar projections of 1979–2013 seasonal trends ( $\text{m yr}^{-1}$ ) in 500 hPa geopotential heights. Trends are computed for winter (a; December, January, February (DJF)), spring (b; March, April, May (MAM)), summer (c; June, July, August (JJA)) and autumn (d; September, October, November (SON)) seasons. Geopotential height fields are sourced from NCEP–DOE–R2.

<sup>1</sup>Department of Earth System Science, Stanford University, Stanford, California 94305, USA. <sup>2</sup>Woods Institute for the Environment, Stanford University, Stanford, California 94305, USA. <sup>3</sup>International Pacific Research Center, University of Hawaii at Manoa, Honolulu, Hawaii 96822, USA. <sup>4</sup>Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA. <sup>5</sup>Cooperative Institute for Climate Science, Princeton University, Princeton, New Jersey 08540, USA. <sup>6</sup>Department of Statistics, Stanford University, Stanford, California 94305, USA.

|                    |     | DJF |    |   |   |   | MAM |    |   |   |   | JJA |    |   |   |   | SON |    |   |   |   |
|--------------------|-----|-----|----|---|---|---|-----|----|---|---|---|-----|----|---|---|---|-----|----|---|---|---|
|                    |     | 05  | 95 | O | P | M | 05  | 95 | O | P | M | 05  | 95 | O | P | M | 05  | 95 | O | P | M |
| Europe             | Sat |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
|                    | Ice | +   |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
| Western Asia       | Sat |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
|                    | Ice |     | +  |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
| Central Asia       | Sat |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
|                    | Ice | +   |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
| Eastern Asia       | Sat |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
|                    | Ice | +   |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
| Western N. America | Sat |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
|                    | Ice |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
| Central N. America | Sat |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
|                    | Ice |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
| Eastern N. America | Sat |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |
|                    | Ice |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |     |    |   |   |   |

|  |                           |
|--|---------------------------|
| 500 hPa circulation pattern occurrence (O), persistence (P) and maximum duration (M) | Robust change             |
| Reanalyses surface temperature extremes Cold (05) and Hot (95)                       | Equivocal                 |
| 1979–2013 (sat) and 1990–2013 (ice) trends   | Cooling                   |
|  | Warming                   |
|  | (+) positive (–) negative |

**Figure 2 | Trends in surface temperature extremes and atmospheric circulation patterns.** Trends are calculated for each Northern Hemisphere season (December, January, February (DJF), winter; March, April, May (MAM), spring; June, July, August (JJA), summer; September, October, November (SON), autumn) for two periods: 1979–2013 (satellite era) and 1990–2013 (ice era). Regional domains (see Fig. 3a) in which one or more of the four SOM circulation patterns demonstrate robust trends in mid-atmospheric circulation pattern occurrence (O), persistence (P), or maximum duration (M) are shown in green (Extended Data Figs 1 and 2). Positive (+) and negative (–) symbols are displayed when all three reanalyses show statistically significant trends in a particular circulation pattern (5% significance level; Methods), and agree on the sign of those trends. Multiple symbols within a box indicate multiple robust pattern trends. White boxes without symbols indicate no statistically significant trends and/or reanalysis disagreement (see Methods). Regional domains with positive and/or negative trends in cold (05) and/or hot (95) extremes receive (+) or (–) symbols when the three reanalyses agree on the sign of the area-weighted trend. Red and blue boxes indicate that the extreme temperature trend results in warming and cooling, respectively, while grey boxes indicate reanalysis disagreement.

matching circulation patterns from all three reanalyses have statistically significant trends of the same sign. We assess the robustness of trends for seven mid-latitude regions (Figs 2 and 3a), over both the era of satellite observation (1979–2013; ‘satellite era’) and the era of rapidly diminishing Arctic sea ice<sup>22</sup> (1990–2013; ‘ice era’). We report circulation patterns that pass these robustness criteria, but also discuss results in the context of (1) comprehensive multiple hypothesis testing; (2) removal of the assumption of linear time-series relationships; (3) use of fewer/more clusters; and (4) addition of atmospheric thermal dilation controls (Methods; Extended Data Table 1 and Extended Data Figs 1–3).

Of the 112 total circulation patterns analysed in each period (Methods), the three reanalyses exhibit statistically significant trends in pattern occurrence for a total of 17, 16 and 16 patterns in the satellite era, and 15, 13 and 14 patterns in the ice era (Extended Data Table 1a). Of these significant occurrence trends, 12 satellite-era and 10 ice-era patterns are robustly significant across all three reanalyses (Fig. 2). The majority of robust satellite-era trends occur in summer and autumn, while robust ice-era trends are more evenly distributed over summer, autumn and winter. These patterns are diverse, and include anticyclonic, cyclonic and ‘dipole’ circulations (Extended Data Figs 4 and 5). Patterns with robust trends in both satellite and ice eras are limited to summer and autumn over western Asia and eastern North America.

While the number of significant trends in pattern persistence varies from 5 to 10 across the individual reanalyses (Extended Data Table 1a),

only three robust pattern persistence trends are identified in each period (Fig. 2). Robust maximum duration trends are more prevalent, including five in the satellite era and six in the ice era. These are predominantly associated with summer anticyclonic patterns, although the maximum duration of central Asia winter troughing events demonstrates a robust ice-era increase (Extended Data Figs 4 and 5). In regions with robust trends in multiple patterns, those patterns are generally complimentary. For example, in summer over eastern North America, robustly increasing satellite-era trends in anticyclonic patterns co-occur with robustly decreasing trends in cyclonic patterns.

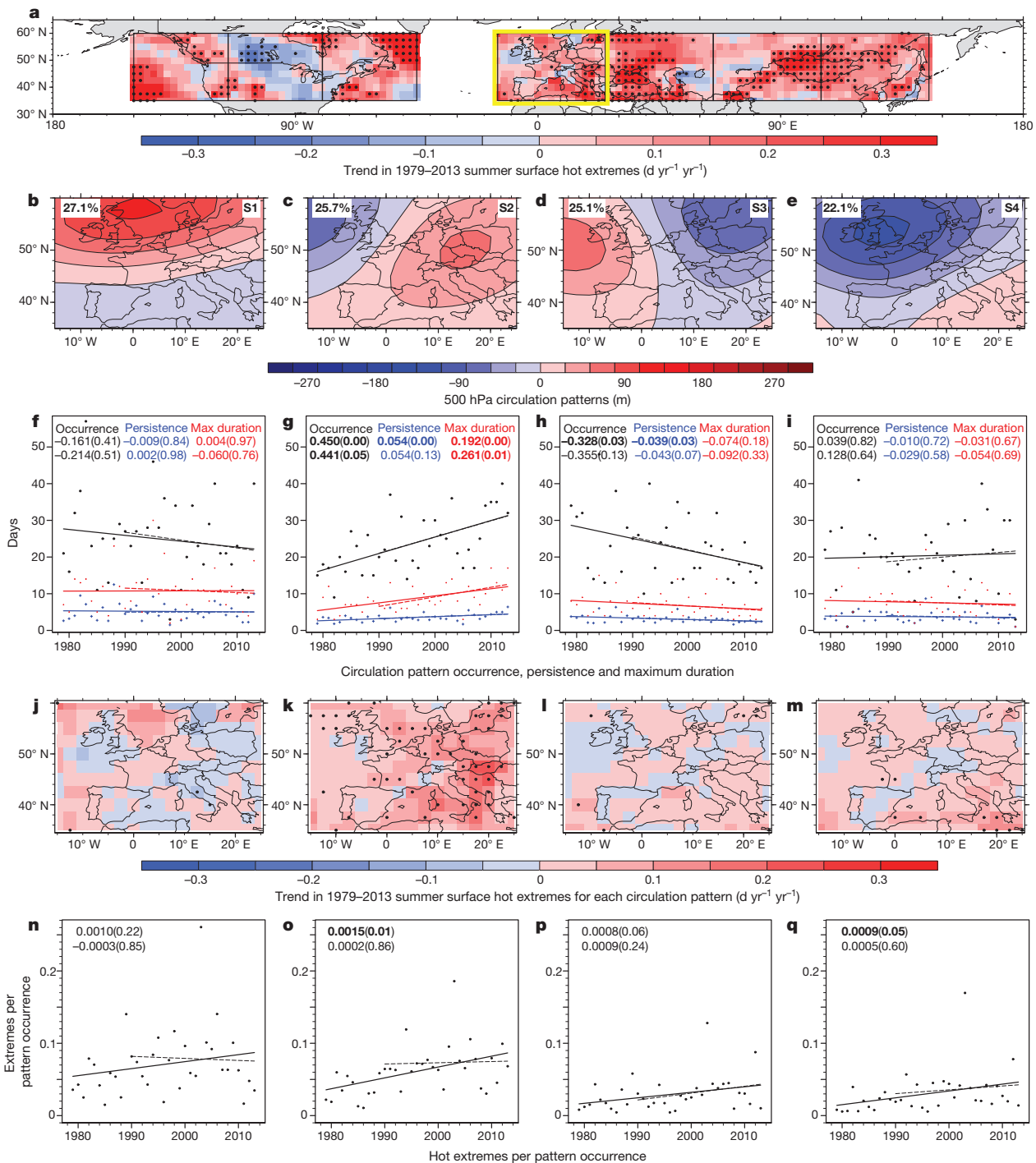
We next explore the extent to which trends in mid-atmospheric circulation patterns have influenced the likelihood of temperature extremes. For each period, we compute area-weighted trends in the seasonal occurrence of temperature extremes for all days, and for those days associated with each SOM pattern (for example, Fig. 3a, j–m; see Methods). The three reanalyses generally agree on the direction of all-days trends: consistent with enhanced radiative forcing and global warming, most regions and seasons show positive trends in hot occurrence, and negative trends in cold occurrence (Fig. 2).

Hot extremes are projected to increase due to the dynamic and thermodynamic effects of global warming<sup>1,23</sup>. Consistent with other assessments<sup>1,2</sup>, we find substantial increases in extreme heat occurrence over the mid-latitudes (Extended Data Figs 6 and 7). For instance, the regional-mean occurrence of summer hot days over Europe, western Asia and eastern North America has increased 0.10, 0.16 and 0.13 d yr<sup>−1</sup> yr<sup>−1</sup>, respectively, over the satellite era (Fig. 3a and Extended Data Table 2a–c). By definition, one would expect (on average) ~4.5 5th/95th percentile events per 3-month season, meaning that an increase of 0.10 d yr<sup>−1</sup> yr<sup>−1</sup> accumulated over the course of the satellite era (35 years) yields an additional ~3.5 d yr<sup>−1</sup>, an ~75% increase.

Heatwaves, similar to those which occurred in western Russia in 2010 and Europe in 2003, develop when persistent anticyclonic patterns, often referred to as ‘atmospheric blocking’, initiate a cascade of self-reinforcing, heat-accumulating physical processes<sup>24,25</sup>. In addition to the increasing trends in extreme heat occurrence, robust positive trends in the occurrence, persistence and maximum duration of satellite-era summer mid-atmospheric anticyclonic patterns are detected over Europe (Fig. 3c, g), western Asia (Fig. 4a, e), and eastern North America (Extended Data Fig. 4c). Robust positive trends in the occurrence of satellite-era anticyclonic patterns are also detected—along with increasing hot extremes—in autumn over eastern North America (Fig. 4c, g), eastern Asia (Fig. 4d, h) and central North America, and in spring over Europe (Extended Data Fig. 1).

Increases in hot extremes may result from dynamic changes (namely greater occurrence and persistence of anticyclonic patterns) as well as from thermodynamic changes such as land cover change or global warming (reflected in the increased intensity of extreme temperature when anticyclonic patterns occur). Over Europe, the summer occurrence of circulations similar to dipole patterns with ridging over the eastern half of the domain (Fig. 3c) increased 0.45 d yr<sup>−1</sup> yr<sup>−1</sup> over the satellite era, while the persistence and maximum duration increased 0.05 and 0.19 d event<sup>−1</sup> yr<sup>−1</sup>, respectively (Fig. 3g). The trend in the frequency of hot events coincident with this pattern (0.06 d yr<sup>−1</sup> yr<sup>−1</sup>; Fig. 3k and Extended Data Table 2a) accounts for ~61.5% of the total trend in hot extremes over Europe (0.10 d yr<sup>−1</sup> yr<sup>−1</sup>; Fig. 3a). In addition, the number of hot extremes per pattern occurrence has increased for all four patterns (Fig. 3n–q and Extended Data Table 2a). Under the assumption of pattern stationarity, we perform a quantitative partitioning of the dynamic and thermodynamic contributions to the overall extreme temperature trend, as well as to the extreme temperature trends associated with each pattern<sup>10</sup> (Methods). This partitioning reveals that ~27.3% of the 0.10 d yr<sup>−1</sup> yr<sup>−1</sup> overall increasing trend in hot extremes is driven by the dynamic influence of increased occurrence of the dipole pattern (along with ~35.2% thermodynamic and ~1.0% interaction influences). Additionally, of the





**Figure 3 | Trends in circulation patterns and hot extremes over Europe.**

**a**, 1979–2013 trends in summer hot extreme occurrences for all regional domains based on 2-m maximum/minimum temperatures from NCEP–DOE–R2. **b–e**, SOM-derived mid-atmospheric circulation patterns (500 hPa geopotential height anomalies) over Europe. White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. **f–i**, Time series of SOM circulation pattern occurrence (black ( $\text{d yr}^{-1}$ )), persistence (blue ( $\text{d event}^{-1}$ )) and maximum duration (red ( $\text{d event}^{-1}$ )). The slope of the trend line ( $\text{yr}^{-1}$ ) and *P* values (in parentheses) are colour coded,

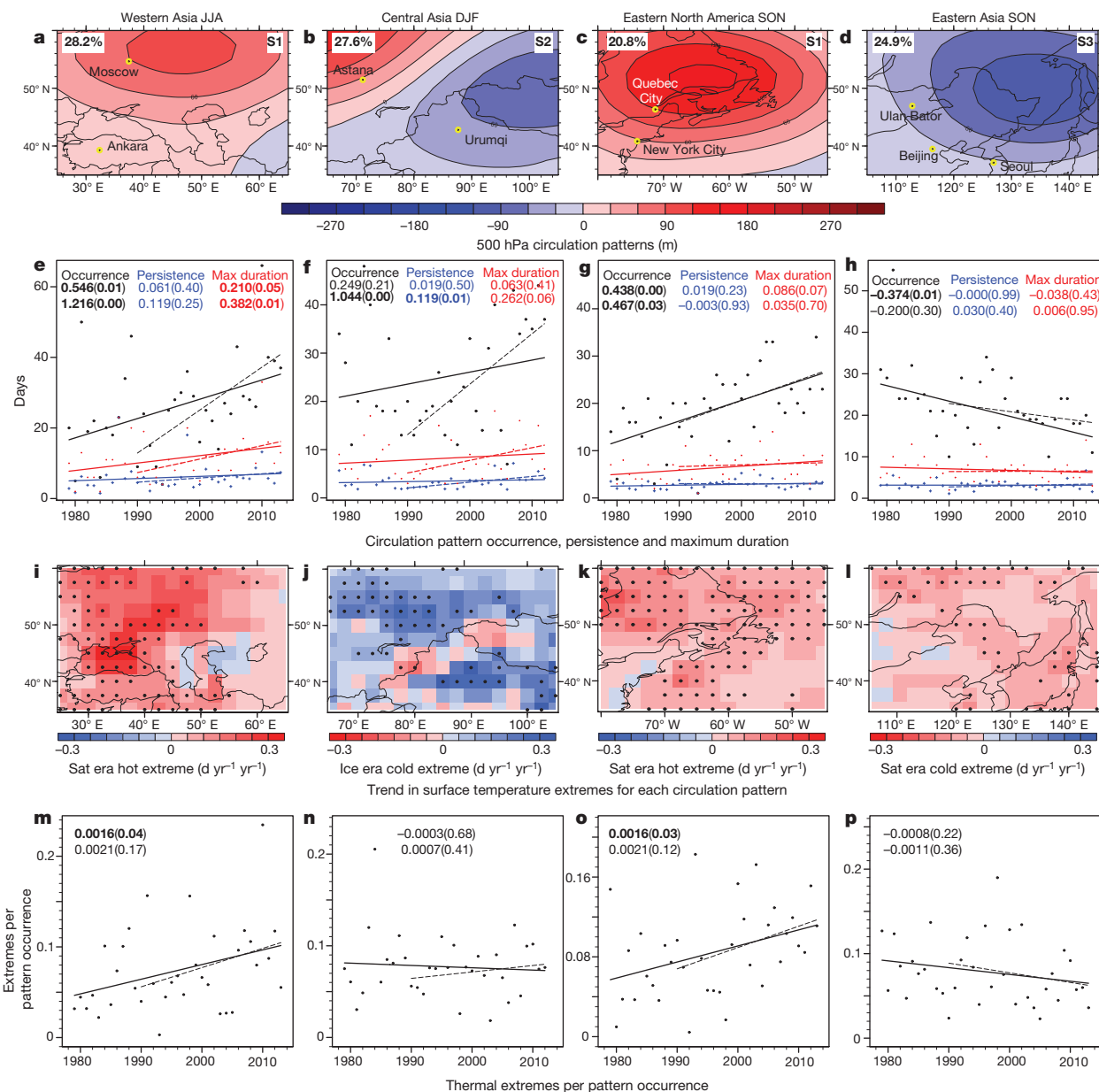
$0.06 \text{ d yr}^{-1} \text{yr}^{-1}$  portion of the hot extremes trend associated with this pattern,  $\sim 57.3\%$  is attributable to thermodynamic influences and  $\sim 44.3\%$  to increased pattern occurrence (Extended Data Table 2a). Together, these results suggest that the observed increase in extreme summer heat over Europe is attributable to both increasing frequency of blocking circulations and changes in the surface energy balance.

with the values from 1979–2013 (solid trend line) displayed above those from 1990–2013 (dashed trend line). **j–m**, Spatially rendered trends in hot extreme occurrences for days that correspond to each SOM circulation pattern.

**n–q**, Time series of the area-weighted mean of hot extremes per pattern occurrence, referred to throughout the text as a measure of the intensity of temperature extremes associated with each pattern. Statistically significant trends (5% significance level; Methods) are shown by stippling in the mapped panels and by bold font in the scatter plots.

Similar results are found in other regions that exhibit robust upward trends in anticyclonic patterns (Fig. 4 and Extended Data Table 2).

Global warming is also generally expected to decrease the frequency of cold extremes<sup>1</sup>. In autumn over eastern Asia, the occurrence of satellite-era cold extremes decreased  $0.08 \text{ d yr}^{-1} \text{yr}^{-1}$  (Extended Data Fig. 6g and Extended Data Table 2f), indicating a reduction of



**Figure 4 | Circulation pattern and thermal extreme trends for selected regions.** Trends in thermal extreme occurrences for selected regions and seasons based on 2-m maximum/minimum temperatures from NCEP-DOE-R2. **a–d**, SOM-derived mid-atmospheric circulation patterns (500 hPa geopotential height anomalies) over western Asia in summer (**a**), central Asia in winter (**b**), eastern North America in autumn (**c**), and eastern Asia in autumn (**d**). White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. In contrast to Fig. 3, just one of the four SOM circulation patterns is displayed from each region. **e–h**, Time series of SOM circulation pattern occurrence (black (d yr<sup>-1</sup>), persistence (blue (d event<sup>-1</sup>)) and maximum duration (red (d event<sup>-1</sup>)). The slope of the trend line (yr<sup>-1</sup>)

~60% over the 35-year period. A majority of this decreasing trend (0.05 d yr<sup>-1</sup> yr<sup>-1</sup>; Extended Data Table 2f) is attributable to changes associated with one pattern type: cyclonic circulations capable of advecting cold air equator-ward (Fig. 4d). Less frequent occurrence of cyclonic patterns (Fig. 4h), in conjunction with less intense cold temperature anomalies when cyclonic patterns occur (Fig. 4p), drives ~62.5% of the overall decreasing trend (partitioned ~35.4% dynamic, ~21.9% thermodynamic, ~5.3% interaction). Of the 0.05 d yr<sup>-1</sup> yr<sup>-1</sup> decrease in extreme cold associated with the trend in cyclonic patterns, partitioning indicates ~56.5% dynamic and ~35.0% thermodynamic influences (Extended Data Table 2f).

and *P* values (in parentheses) are colour coded, with the values from 1979 to 2013 (solid trend line) displayed above those from 1990 to 2013 (dashed trend line). **i–l**, Spatially rendered trends in thermal extreme occurrences for days that correspond to each SOM circulation pattern. **m–p**, Time series of the area-weighted mean of temperature extremes per pattern occurrence, referred to throughout the text as a measure of the intensity of temperature extremes associated with each pattern. Statistically significant trends (5% significance level; Methods) are shown by stippling in the mapped panels and by bold font in the scatter plots. Refer to Extended Data Figs 6 and 7 for satellite-era and ice-era trends in temperature extremes over the regional domains.

In contrast to this expected extreme cold decrease, winter cold extremes over central Asia have increased 0.07 d yr<sup>-1</sup> yr<sup>-1</sup> over the ice era (Extended Data Fig. 7a). One-hundred-and-forty-nine per cent of this trend (0.10 d yr<sup>-1</sup> yr<sup>-1</sup>; Fig. 4j) occurred when mid-atmospheric circulation was similar to a pattern of troughing in the south and east, and ridging in the northwest (Fig. 4b; ~111.4% dynamic, ~25.6% thermodynamic, ~11.6% interaction). Trend percentages exceeding 100% indicate that other circulation patterns provide negative contributions. Occurrence and persistence of this dipole pattern robustly increased (1.0 d yr<sup>-1</sup> yr<sup>-1</sup> and 0.12 d event<sup>-1</sup> yr<sup>-1</sup>, Fig. 4f) at the expense of all other circulations (Extended Data Table 2d).

Partitioning indicates that ~75.0% of the extreme cold trend associated with this pattern is due to the dynamic influence of increased pattern occurrence, with ~17.2% linked to thermodynamic influences (Extended Data Table 2d).

Substantial dynamic contributions to the overall trend in cold extremes could be expected given that circulations that support the equator-ward advection of Arctic air will bring anomalously cold temperatures to lower-latitude locales<sup>20</sup>. Increased occurrence of such patterns has previously been observed, and linked to reduced regional sea-ice and decreased baroclinicity over the Barents and Kara seas<sup>4,17,26–28</sup>. Positive thermodynamic contributions to the extreme cold trend indicate processes that are in opposition to the direct warming effects of enhanced radiative forcing. For example, positive thermodynamic contributions from three of the four winter patterns over central Asia (Extended Data Table 2d) suggest that these thermodynamic contributions are largely independent of atmospheric circulation, and therefore potentially related to surface processes such as increased snow cover and enhanced diabatic cooling<sup>4,28</sup>.

The circulation trends detected here cannot as yet be attributed to anthropogenic or natural causes, nor can they be projected to continue into the future. Attribution and projection will require an increased understanding of the causes of the circulation trends, including the ability to identify the signal of an anthropogenically forced trend from the noise of internal decadal-scale climate variability<sup>16,29</sup>. However, our quantitative partitioning, in conjunction with targeted climate model simulations<sup>16,29,30</sup>, offers the potential to fingerprint dynamic and thermodynamic climate influences in isolation, which in turn may facilitate attribution of the observed trends, and projection of future trends. We hypothesize that the main assumption of our quantitative partitioning—pattern stationarity—is justified given the expectation that circulation responses to enhanced radiative forcing are likely to reinforce pre-existing modes of natural variability<sup>15,16</sup>. A related assumption is that the reanalyses act as reasonable proxies for the state of the three-dimensional atmosphere through time. Given uncertainties in the data assimilation and numerical modelling that underpin atmospheric reanalysis, we have restricted our identification criteria to those trends that are statistically significant in all three reanalyses.

Our approach finds robust trends in mid-atmospheric circulation patterns over some regions, and suggests that both dynamic and thermodynamic effects have contributed to observed changes in temperature extremes over the past 35 years. Although thermodynamic influences have largely dominated these changes, dynamic influences have been critical in some regions and seasons. Long-term projections of future dynamic contributions are challenging given the substantial underlying decadal-scale variability, as well as the uncertain impact of anthropogenic forcing on mid-latitude circulation<sup>15,16</sup>. However, given our finding that many patterns have exhibited increasing (decreasing) intensity of extreme hot (cold) events, and that those trends are coincident with a nearly categorical increase in thermodynamic forcing, the observed trends of increasing hot extremes and decreasing cold extremes could be expected to continue in the coming decades, should greenhouse gases continue to accumulate in the atmosphere.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 6 November 2014; accepted 27 April 2015.**

- Field, C. B. *et al.* (eds) *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (Cambridge Univ. Press, 2012).
- Donat, M. G. *et al.* Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *J. Geophys. Res.* **118**, 2098–2118 (2013).
- Francis, J. A. & Vavrus, S. J. Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophys. Res. Lett.* **39**, L06801 (2012).
- Liu, J., Curry, J. A., Wang, H., Song, M. & Horton, R. M. Impact of declining Arctic sea ice on winter snowfall. *Proc. Natl Acad. Sci. USA* **109**, 4074–4079 (2012).

- Petoukhov, V., Rahmstorf, S., Petri, S. & Schellnhuber, H. J. Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes. *Proc. Natl Acad. Sci. USA* **110**, 5336–5341 (2013).
- Screen, J. A. & Simmonds, I. Caution needed when linking weather extremes to amplified planetary waves. *Proc. Natl Acad. Sci. USA* **110**, E2327 (2013).
- Barnes, E. A., Dunn-Sigouin, E., Masato, G. & Woolings, T. Exploring recent trends in Northern Hemisphere blocking. *Geophys. Res. Lett.* **41**, 638–644 (2014).
- Screen, J. & Simmonds, I. Exploring links between Arctic amplification and mid-latitude weather. *Geophys. Res. Lett.* **40**, 959–964 (2013).
- Kohonen, T. *Self-Organizing Maps* 501 (Springer, 2001).
- Cassano, J. J., Uotila, P., Lynch, A. H. & Cassano, E. N. Predicted changes in synoptic forcing of net precipitation in large Arctic river basins during the 21<sup>st</sup> century. *J. Geophys. Res.* **112**, G04S49 (2007).
- Johnson, N. C., Feldstein, S. B. & Trenberth, B. The continuum of Northern Hemisphere teleconnection patterns and a description of the NAO shift with the use of self-organizing maps. *J. Clim.* **21**, 6354–6371 (2008).
- Lee, S. & Feldstein, S. B. Detecting ozone- and greenhouse gas-driven wind trends with observational data. *Science* **339**, 563–567 (2013).
- Diffenbaugh, N. S. *et al.* in *Climate Change 2014: Impacts, Adaptation, and Vulnerability* (eds Field, C. B. *et al.*) 137–141 (IPCC, Cambridge Univ. Press, 2014).
- Field, C. B. *et al.* in *Climate Change 2014: Impacts, Adaptation, and Vulnerability* (eds Field, C. B. *et al.*) 1–32 (IPCC, Cambridge Univ. Press, 2014).
- Shepherd, T. G. Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.* **7**, 703–708 (2014).
- Deser, C., Phillips, A. S., Alexander, M. A. & Smoliak, B. V. Projecting North American climate over the next 50 years: uncertainty due to internal variability. *J. Clim.* **27**, 2271–2296 (2014).
- Cohen, J. *et al.* Recent Arctic amplification and extreme mid-latitude weather. *Nat. Geosci.* **7**, 627–637 (2014).
- Palmer, T. Record-breaking winters and global climate change. *Science* **344**, 803–804 (2014).
- Wallace, J. M., Held, I. M., Thompson, D. W. J., Trenberth, K. E. & Walsh, J. E. Global warming and winter weather. *Science* **343**, 729–730 (2014).
- Screen, J. A. Arctic amplification decreases temperature variance in northern mid-to high-latitudes. *Nature Clim. Change* **4**, 577–582 (2014).
- Hartmann, D. L. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) 159–254 (IPCC, Cambridge Univ. Press, 2013).
- Simmonds, I. Comparing and contrasting the behaviour of Arctic and Antarctic sea ice over the 35-year period 1979–2013. *Ann. Glaciol.* **56**, 18–28 (2015).
- Diffenbaugh, N. S. & Ashfaq, M. Intensification of hot extremes in the United States. *Geophys. Res. Lett.* **37**, L15701 (2010).
- Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C. & Vila-Guerau de Arellano, J. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.* **7**, 345–349 (2014).
- Diffenbaugh, N. S., Pal, J. S., Trapp, R. J. & Giorgi, F. Fine-scale processes regulate the response of extreme events to global climate change. *Proc. Natl Acad. Sci. USA* **102**, 15774–15778 (2005).
- Inoue, J., Hori, M. E. & Takaya, K. The role of Barents Sea ice in the wintertime cyclone-track and emergence of a warm-Arctic cold-Siberian Anomaly. *J. Clim.* **25**, 2561–2568 (2012).
- Mori, M., Watanabe, M., Shiogama, H., Inoue, J. & Kimoto, M. Robust Arctic sea-ice influence on the frequent Eurasian cold winters in past decades. *Nat. Geosci.* **7**, 869–873 (2014).
- Cohen, J., Furtado, J., Barlow, J. M., Alexeev, V. & Cherry, J. Arctic warming, increasing fall snow cover and widespread boreal winter cooling. *Environ. Res. Lett.* **7**, 014007 (2012).
- Screen, J. A., Deser, C., Simmonds, I. & Tomas, R. Atmospheric impacts of Arctic sea-ice loss, 1979–2009: Separating forced change from atmospheric internal variability. *Clim. Dyn.* **43**, 333–344 (2014).
- Screen, J. A., Simmonds, I., Deser, C. & Tomas, R. The atmospheric response to three decades of observed Arctic sea ice loss. *J. Clim.* **26**, 1230–1248 (2013).

**Acknowledgements** Work by D.E.H., D.S., D.L.S. and N.S.D. was supported by NSF CAREER Award 0955283, DOE Integrated Assessment Research Program Grant No. DE-SC005171DE-SC005171, and a G.J. Lieberman Fellowship to D.S. Contributions from N.C.J. were supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections program award NA14OAR4310189. B.R. acknowledges support from the US Air Force Office of Scientific Research (FA9550-13-1-0043), the US National Science Foundation (DMS-0906392, DMS-CMG-1025465, AGS-1003823, DMS-1106642, and DMS-CAREER-1352656), the Defense Advanced Research Projects Agency (DARPA YFA N66001-111-4131), and the UPS Foundation (SMC-DBNKY). We thank B. Santer, J. Cattiaux, D. Touma, and J. S. Mankin for discussions that improved the manuscript. Computational resources for data processing and analysis were provided by the Center for Computational Earth and Environmental Science in the School of Earth, Energy, and Environmental Sciences at Stanford University.

**Author Contributions** D.E.H. conceived the study. D.E.H., N.C.J., D.S., D.L.S. and N.S.D. designed the analysis and co-wrote the manuscript. D.E.H., N.C.J. and D.S. provided analysis tools. D.E.H. performed the analysis. B.R. provided and described the multiple hypothesis testing and transformation analysis.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.E.H. ([danethan@stanford.edu](mailto:danethan@stanford.edu)).



## METHODS

**Categorization of circulation patterns.** We use SOM cluster analysis<sup>9–12</sup> to categorize large-scale circulation patterns over seven Northern Hemisphere domains<sup>31</sup> using daily 500 hPa geopotential height anomaly fields from the NCAR/NCEP-R1 (ref. 32), NCEP-DOE-R2 (ref. 33) and ECMWF ERA-Interim (ref. 34) reanalyses. Daily anomalies are calculated by subtracting the seasonal cycle (calendar-day mean) from each grid cell. Reanalyses are analysed individually to maintain their physical consistency, and to facilitate their intercomparison. The SOM's unsupervised learning algorithm requires neither a priori knowledge of which types of circulation patterns might be detected, nor the specific geographic regions in which they might occur. Geopotential height anomaly fields from each day are assigned to one of a pre-defined number of nodes, according to pattern similarity. The final SOM patterns are obtained by minimizing the Euclidean distance between iteratively updated nodes and their matching daily geopotential height anomaly fields<sup>11</sup>. Each SOM pattern can therefore be viewed as a representative composite of relatively similar circulation patterns.

Owing to global-scale warming, trends in geopotential height anomalies record both altered atmospheric circulation patterns and the thermal expansion of the troposphere. To isolate the signal of circulation pattern change, previous clustering analyses have assumed uniform thermal dilation and removed either the domain average<sup>35</sup> or domain average linear trend<sup>36</sup> from the daily-scale anomalies. In our analysis, we find that 1979–2013 trends in Northern Hemisphere geopotential height anomalies are non-uniform in both magnitude and sign (Fig. 1), and demonstrate substantial seasonal, regional and latitudinal differences (Extended Data Fig. 3e). These findings suggest that for the relatively short period of our analysis, an assumption of uniform thermal dilation is inappropriate. Moreover, the strong spatial heterogeneity indicates the importance of large-scale dynamics in the regional geopotential height trends, and so the removal of local geopotential height trends would conflate dynamic changes with thermal dilation. Therefore, in the main text we present results and conclusions based on raw geopotential height data. However, despite the lack of uniform expansion, we have conducted an analysis that attempts to account for the effects of thermal dilation. SOM analyses are performed on geopotential heights that have been detrended by removing the seasonal mean hemispheric trend (Extended Data Fig. 3e) from each grid cell. Results from this analysis indicate that the magnitude, significance and sign of circulation trends are sensitive to the method of controlling for thermal expansion (Extended Data Fig. 3f–j). Despite this sensitivity, the conclusions presented in the main text are supported, in that both raw and detrended analyses generally suggest trends of similar magnitude, sign and significance.

Based on domain-wide pattern correlations between daily height field anomalies and different SOM node counts<sup>12</sup> (Extended Data Fig. 8), we divide circulation patterns over each domain into four SOM nodes (for example, Fig. 3b–e). To determine a suitable number of nodes, a suite of different node counts were analysed. We found that four nodes were sufficiently great in number to capture a diversity of highly generalized circulation patterns, but sufficiently few to facilitate convenient presentation and, critically, to prevent overly similar SOM patterns<sup>12</sup>. To test the sensitivity of our results to the number of nodes, we present 2-, 4-, 8- and 16-node SOMs for the summer season over the European domain (Extended Data Figs 1 and 2). Based on these analyses, it is apparent that a 2-node SOM is insufficient to capture the diversity of circulation patterns that are found in the reanalysis data (Extended Data Fig. 1a), whereas 8- and 16-node SOMs produce nodes with overly similar circulation patterns (Extended Data Figs 1c and 2). Examination of the 4-, 8- and 16-node SOMs largely verifies the conclusions drawn from the 4-node SOM: that the occurrence of patterns with ridging over the eastern half of the European domain has increased over time, while the occurrence of complimentary patterns has decreased. We note that these pattern trends are not identified in the 2-node SOM, confirming that two nodes are too few to capture specific circulation patterns that are critical for extreme temperature occurrence. Similar node-count analyses for other regions/seasons likewise verify the conclusions drawn from the 4-node analyses that are presented in the main text (not shown).

**Calculation of robust trends in circulation patterns.** For each season in each year, we calculate (1) the total number of days on which each SOM pattern occurs (occurrence ( $\text{d yr}^{-1}$ )); (2) the mean length of consecutive occurrence (persistence ( $\text{d event}^{-1}$ )); and (3) the longest consecutive occurrence (maximum duration ( $\text{d event}^{-1}$ )). A trend in one of these characteristics is considered robust when the trend in that pattern is (1) statistically significant in all three reanalyses, and (2) of the same sign in all three reanalyses. Trends with regression coefficients that surpass the 5% significance (95% confidence) threshold are considered statistically significant. Trends are calculated across satellite-era and ice-era annual time series using the approach of ref. 37, which allows us to account for temporal dependence. Here, the trends are calculated using linear least squares regression, but to account for temporal dependence, the confidence bounds of annual time-series trends with

lag-one autocorrelation greater than the 5% significance level are recalculated by adjusting the number of degrees of freedom used to compute the regression coefficient significance<sup>37</sup>. As a result, unlike a simple linear regression, this approach does not rely on the independent and identically distributed assumption for the residuals, but instead accounts for temporal dependence using a red noise assumption.

The approach of ref. 37 also assumes that the distribution of the residuals is Gaussian. Using the Anderson–Darling<sup>38</sup> test for normality, we find that 91–100% of residual distributions in each metric of each reanalysis do not reject the null hypothesis of Gaussianity when multiple hypothesis testing controlling the family-wise error rate<sup>39</sup> (FWER) at the 5% significance level is considered (Extended Data Table 1b). The Gaussianity assumption is therefore largely appropriate. However, due to the identification of non-normality in some distributions, particularly in the persistence and maximum duration metrics, we apply Box–Cox power transformations<sup>40</sup> to all distributions. Using the Anderson–Darling test, we find that 96–100% of the distributions of the residuals in the transformed setting do not reject the null hypothesis of Gaussianity at the 5% significance level. Furthermore, when multiple hypothesis testing is considered by controlling the FWER at the 5% level, 100% of the individual tests are non-significant (Extended Data Table 1b). In addition, the number of pattern trends identified as significant in the transformed case is largely consistent with the non-transformed regression analysis. However, additional significant trends in the persistence and maximum duration metrics are identified when Box–Cox transformations are used (Extended Data Table 1a). The large overlap of results between the transformed and non-transformed analyses suggests that although individual residual distributions may vary, the Gaussian assumption applied throughout this study is for the most part quite robust, although in some cases not valid. In sum, the non-transformed analysis that fits a linear relationship to circulation pattern metrics allows for a relatively simple classification of two short analysis periods, while simultaneously accounting for temporal dependence in a large number ( $>2,000$ ) of individual time series (7 regions  $\times$  4 nodes  $\times$  4 seasons  $\times$  3 characteristics  $\times$  2 time periods  $\times$  3 reanalyses).

Because SOM nodes are calculated independently for each reanalysis, individual SOM patterns must be matched between the three reanalyses in order to determine whether an individual pattern shows robust results across all three reanalyses. To determine which SOM patterns are the closest match between the three reanalyses, the root mean square error (RMSE) is calculated between the SOM patterns of one reanalysis and those of the other reanalyses. Patterns with the smallest RMSEs are considered matches. Although we undertake a multi-reanalysis robustness evaluation, further work is needed to confirm that other available reanalyses (such as CFSR and MERRA) show the same trends.

**Multiple hypothesis testing of linear trends.** It is possible that some of the trends identified as significant in any individual reanalysis could occur by chance. In addition to screening for those patterns that are significant in all three reanalyses (our ‘robustness’ criterion), we also employ formal multi-hypothesis testing using several methodologies. The first is the familywise error rate (FWER). This type of error metric controls the probability of falsely rejecting any null hypothesis, and is considered one of the strictest forms of error control<sup>39</sup>. Since a certain number of false rejections can happen by chance alone, one can account for this formally by using the  $k$ -familywise error rate<sup>41</sup> ( $k$ -FWER). The  $k$ -FWER controls the probability of falsely rejecting  $k$  or more null hypotheses, and aims to formalize the concept that some of the hypotheses will be rejected by chance. One option for the value of  $k$  is to use the expected number of hypotheses that will be rejected at a given significance level. For instance, in our study, out of 112 total ‘local’ hypotheses, 5 or 6 hypotheses will be significant at the 5% significance level by chance ( $112 \times 0.05 = 5.6$ ). In this case, one can evaluate the probability that 7 or more hypotheses are falsely rejected, since on average about 6 could be rejected as significant by chance. The third metric is the false discovery rate (FDR), which controls the expectation of the ratio given by the number of false rejections divided by the total number of rejections<sup>39</sup>.

All of the above measures of error control aim to guard against hypotheses being falsely declared as significant in the context of multiple tests. To be thorough, we have implemented all three types of error control at both global significance levels of 5% and 10%. The results of these analyses are summarized in Extended Data Table 2. We note that all three metrics heavily favour the null as they are designed to protect against the possibility of false positives. Despite this, the presence of local tests that reject the null represents a strong confirmation of the significance of those local tests. The fact that a number of local hypotheses still prevail as significant, even after imposing much stricter multiple testing error controls, arises partly from the fact that some of the local  $P$  values indicate trends that are so highly significant that they can withstand the stricter multiple testing error control metrics. This rigorous multiple-testing error control yields increased credibility to the scientific conclusions of robust trends in pattern occurrence.



**Temperature extremes.** Daily-scale hot and cold extreme occurrences are calculated using temperature anomalies at each grid cell. Temperature anomalies are computed by removing the seasonal cycle from daily reanalysis 2-m maximum/minimum temperatures. Similar to previous studies<sup>1,2</sup>, temperature extremes are calculated based on the statistical distribution of daily temperature anomalies<sup>20</sup>. Hot/cold extreme thresholds are defined as the 95<sup>th</sup>/5<sup>th</sup> percentile value of the 1979–2013 daily 2-m maximum/minimum temperature anomaly distribution (for example, for JJA, the population of daily-maximum temperature anomalies from the months of June, July and August in the years 1979–2013). Hot/cold extreme occurrences are defined as days on which the daily temperature anomalies are greater/less than (or equal to) the hot/cold extreme thresholds. Reanalysis temperature extremes are qualitatively similar to those found in station-based observations<sup>2</sup>. Given this similarity, we use the reanalysis temperatures in order to maintain internal physical consistency between daily 2-m temperatures and daily atmospheric circulation (as represented by the 500 hPa SOM circulation patterns). Trends in temperature extreme occurrence are computed across satellite-era and ice-era annual time series following the methodology of ref. 37.

**Quantitative partitioning.** To determine the dynamic and thermodynamic contributions to trends in temperature extreme occurrence, we adapt the climate change partitioning methodology of Cassano *et al.*<sup>10</sup>. Our adapted methodology partitions the contributions of dynamic and thermodynamic changes to (1) the overall trend in temperature extreme occurrence and (2) the trends associated with individual SOM circulation patterns. Previous applications of the Cassano *et al.*<sup>10</sup> methodology indicate that partitioning is largely insensitive to the number of SOM nodes used in the analysis<sup>42</sup>. All trends in temperature extremes in the below methodology are area-weighted averages. Following Cassano *et al.*<sup>10</sup>:

$$E = \sum_{i=1}^K E_i f_i$$

where  $E$  is the frequency of extreme temperature occurrence,  $f_i$  is the frequency of occurrence of SOM pattern  $i$ ,  $E_i$  is the frequency of extreme temperature occurrence when SOM pattern  $i$  occurs, and  $K$  is the total number of SOM nodes. We decompose  $E$  and  $f$  into time mean and deviation from time mean components:

$$E = \sum_{i=1}^K (\bar{E}_i + E'_i)(\bar{f}_i + f'_i)$$

Now we differentiate the above equation with respect to time, noting that the mean values are constants:

$$\frac{dE}{dt} = \sum_{i=1}^K \left( \bar{f}_i \frac{dE'_i}{dt} + \bar{E}_i \frac{df'_i}{dt} + \frac{d}{dt}(E'_i f'_i) \right)$$

The derivative on the left-hand side provides the area-weighted average trend in the seasonal occurrence of temperature extremes for all days. The summation on the right-hand side, from left to right, provides the thermodynamic, dynamic and interaction contributions for days associated with each SOM pattern,  $i$ .

The thermodynamic contribution of each circulation pattern's extreme temperature trend assumes that each SOM pattern is stationary in time, and that trends in extremes that result during this pattern are the result of influences unrelated to circulation, such as changes in long-wave radiation from increasing

greenhouse gas concentrations, or changes in surface fluxes of moisture and/or radiation resulting from changes in land cover. The thermodynamic contribution associated with each circulation pattern is determined by taking the product of the trend in the intensity of temperature extremes and the mean occurrence of the circulation pattern. Trends in the intensity of temperature extremes are computed by calculating the trend in area-weighted extreme occurrence per pattern occurrence (for example, Fig. 3n–q).

The dynamic contribution of each circulation pattern's extreme temperature trend assumes that, on average, a circulation pattern is associated with a portion of the total extreme event trend, and that changes in the occurrence frequency of that circulation pattern will modify the occurrence frequency of extreme events. The dynamic contribution associated with each circulation pattern is determined by taking the product of the trend in circulation pattern occurrences and the mean number of extreme events per pattern occurrence.

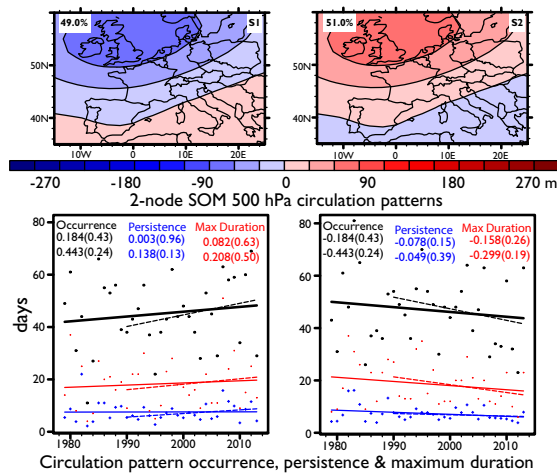
The third component represents the interaction between dynamic and thermodynamic changes, and captures contributions that result from changes in the dynamic component acting on changes in the thermodynamic component, such as the positive/negative feedbacks of surface–atmosphere interactions. The interactive term is determined by computing the trend in the product of circulation pattern occurrence deviations and intensity of temperature extreme deviations.

**Code availability.** SOM code is available at <http://www.cis.hut.fi/projects/somtoolbox/>. All other analysis code is available upon request from the corresponding author (danethan@stanford.edu).

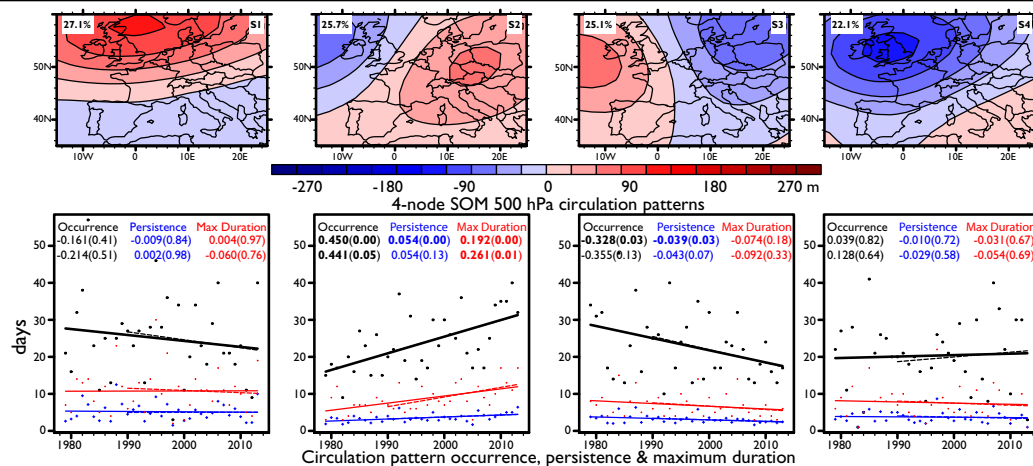
**Reanalysis data sets.** NCAR/NCEP-Reanalysis 1 data was downloaded from <http://www.esrl.noaa.gov/psd/>; NCEP-DOE-Reanalysis 2 data was downloaded from <http://www.esrl.noaa.gov/psd/>; ECMWF ERA-Interim data was downloaded from <http://www.ecmwf.int/>.

31. Screen, J. A. & Simmonds, I. Amplified mid-latitude planetary waves favour particular regional weather extremes. *Nature Clim. Change* **4**, 704–709 (2014).
32. Kalnay, E. *et al.* The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
33. Kanamitsu, M. *et al.* NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Am. Meteorol. Soc.* **83**, 1631–1643 (2002).
34. Dee, D. P. *et al.* The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).
35. Driouech, F., Déqué, M. & Sánchez-Gómez, E. Weather regimes–Moroccan precipitation link in a regional climate change simulation. *Glob. Plan. Ch.* **72**, 1–10 (2010).
36. Cattiaux, J., Douville, H. & Peings, P. European temperatures in CMIP5: origins of present-day biases and future uncertainties. *Clim. Dyn.* **41**, 2889–2907 (2013).
37. Zwiers, F. W. & von Storch, H. Taking serial correlation into account in tests of the mean. *J. Clim.* **8**, 336–351 (1995).
38. Anderson, T. W. & Darling, D. A. Asymptotic theory of certain “goodness-of-fit” criteria based on stochastic processes. *Ann. Math. Stat.* **23**, 193–212 (1952).
39. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc., B* **57**, 289–300 (1995).
40. Box, G. E. P. & Cox, D. R. An analysis of transformations. *J. R. Stat. Soc. B* **26**, 211–243 (1964).
41. Lehmann, E. & Romano, J. P. Generalizations of the familywise error rate. *Ann. Stat.* **33**, 1138–1154 (2005).
42. Skific, N., Francis, J. A. & Cassano, J. J. Attribution of projected changes in atmospheric moisture transport in the Arctic: A self-organizing map perspective. *J. Clim.* **22**, 4135–4153 (2009).

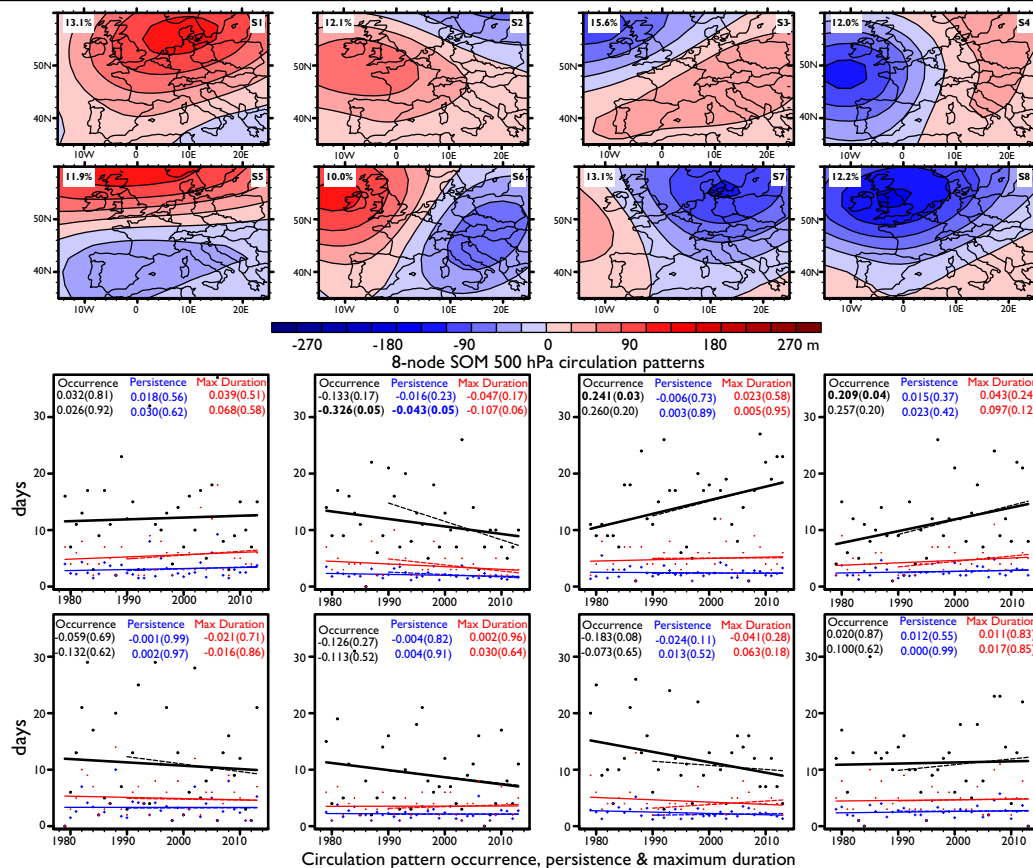
a



b

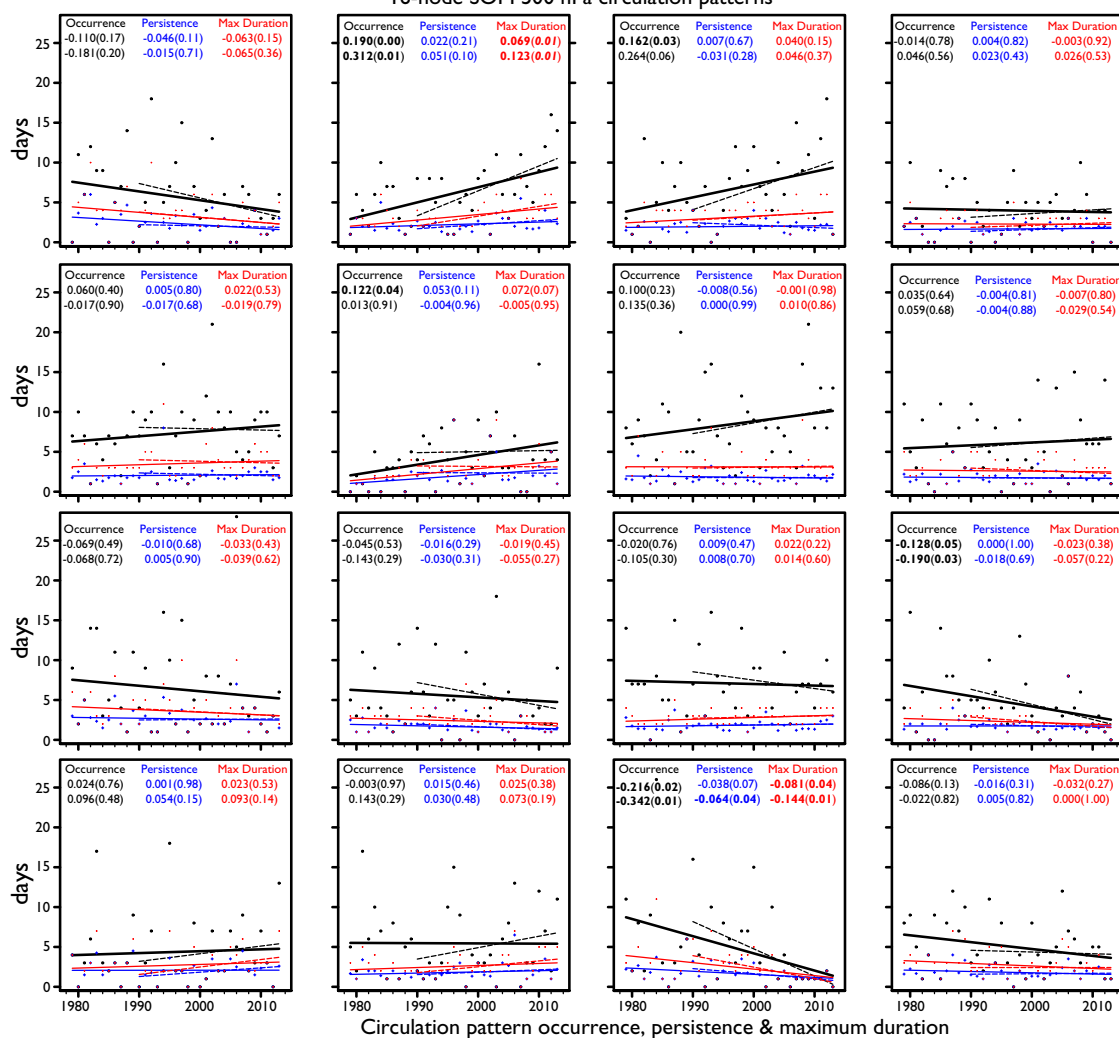
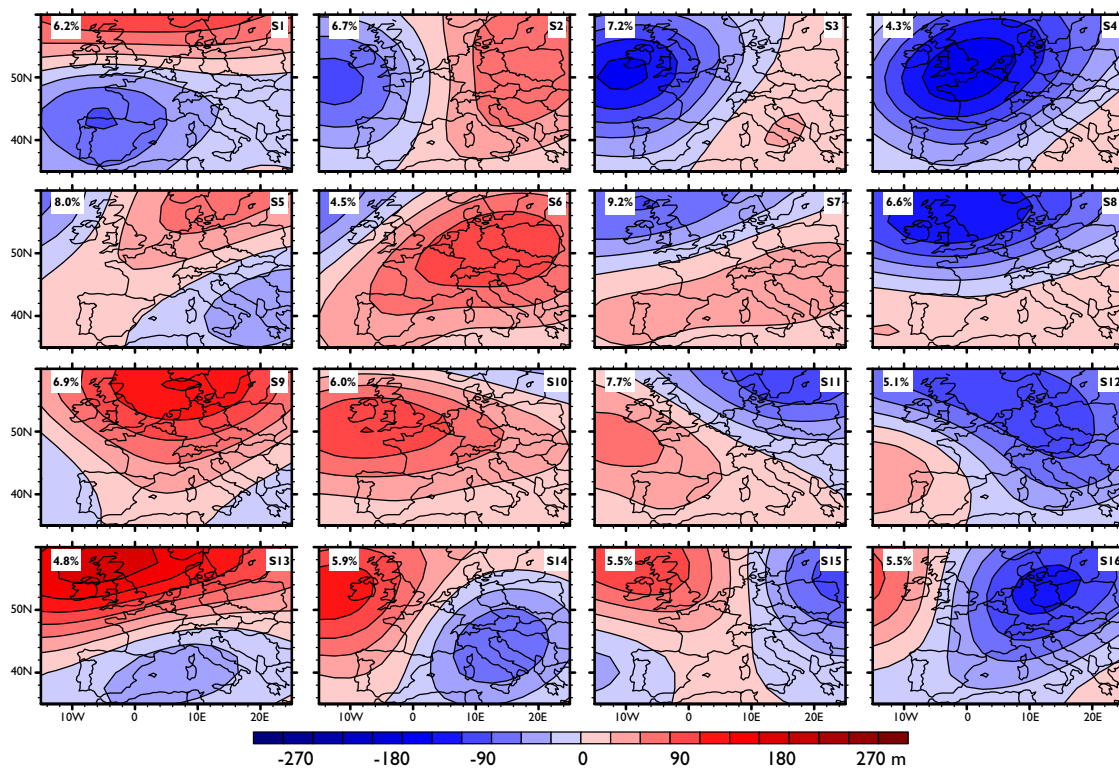


c



**Extended Data Figure 1 | 2-, 4- and 8-node SOM analyses.** SOM-derived mid-atmospheric summer (JJA) circulation patterns (500 hPa geopotential height anomalies) over Europe using 2- (a), 4- (b) and 8-node (c) analyses. White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. Time series of SOM circulation pattern occurrence (black ( $\text{d yr}^{-1}$ )), persistence (blue ( $\text{d event}^{-1}$ )) and maximum duration (red

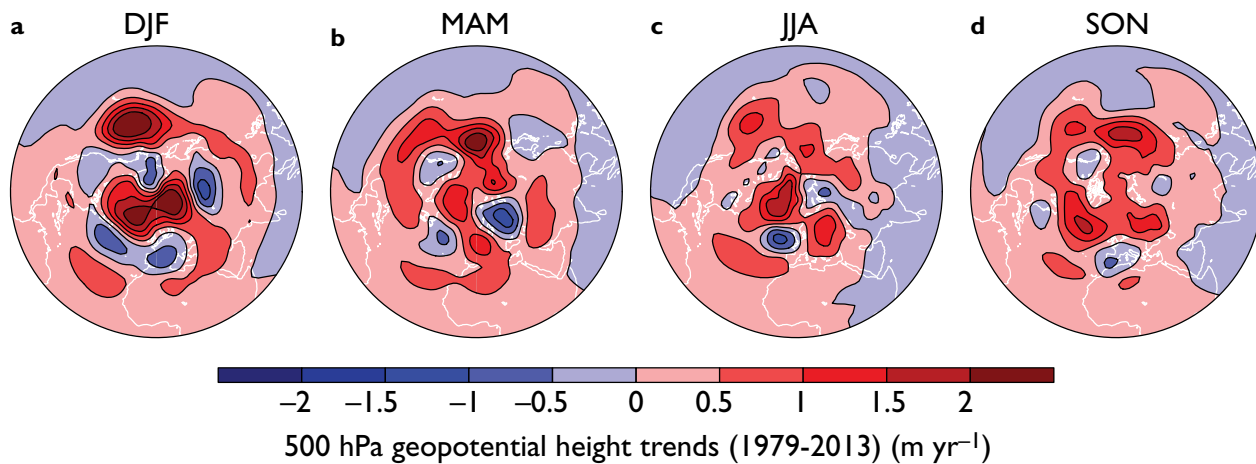
( $\text{d event}^{-1}$ )). The slope of the trend line ( $\text{yr}^{-1}$ ) and  $P$  values (in parentheses) are colour coded, with the values from 1979 to 2013 (solid trend line) displayed above those from 1990 to 2013 (dashed trend line). Statistically significant trends (5% significance level; Methods) are shown by bold fonts in the scatter plots. Geopotential height fields are sourced from the NCEP-DOE-R2 reanalysis<sup>33</sup>.





**Extended Data Figure 2 | 16-node SOM analysis.** SOM-derived mid-atmospheric summer (JJA) circulation patterns (500 hPa geopotential height anomalies) over Europe derived from a 16-node analysis. White boxed values show pattern frequencies in the top left and SOM node numbers in the top right. Time series of SOM circulation pattern occurrence (black ( $\text{d yr}^{-1}$ )), persistence (blue ( $\text{d event}^{-1}$ )) and maximum duration (red ( $\text{d event}^{-1}$ )).

The slope of the trend line ( $\text{yr}^{-1}$ ) and  $P$  values (in parentheses) are colour coded, with the values from 1979 to 2013 (solid trend line) displayed above those from 1990 to 2013 (dashed trend line). Statistically significant trends (5% significance level; Methods) are shown by bold fonts in the scatter plots. Geopotential height fields are sourced from the NCEP-DOE-R2 reanalysis<sup>33</sup>.



e

| 1979-2013 Trends in area-weighted geopotential heights ( $\text{m yr}^{-1}$ ) |        |       |       |       |
|---|--------|-------|-------|-------|
|   | DJF    | MAM   | JJA   | SON   |
| N.Hemisphere  | 0.347  | 0.304 | 0.232 | 0.293 |
| Euro  | -0.210 | 0.708 | 0.221 | 0.053 |
| wAsia   | 0.570  | 0.361 | 0.656 | 0.585 |
| cAsia   | -0.271 | 0.491 | 0.130 | 0.151 |
| eAsia   | 0.005  | 0.100 | 0.439 | 0.551 |
| wNA   | 0.370  | 0.314 | 0.438 | 0.386 |
| cNA   | 0.188  | 0.409 | 0.207 | 0.550 |
| eNA   | 0.370  | 0.319 | 0.527 | 0.818 |

f

| Trend Removed | Euro-JJA-1979-2013 |        |        |      |       |       |        |        |        |       |        |        |
|---------------|--------------------|--------|--------|------|-------|-------|--------|--------|--------|-------|--------|--------|
|               | A                  |        |        | C-A  |       |       | A-C    |        |        | C     |        |        |
|               | OCC                | PER    | DUR    | OCC  | PER   | DUR   | OCC    | PER    | DUR    | OCC   | PER    | DUR    |
| None          | -0.161             | -0.009 | 0.004  | 0.45 | 0.054 | 0.192 | -0.328 | -0.039 | -0.074 | 0.039 | -0.01  | -0.031 |
| N.Hemisphere  | -0.196             | -0.013 | -0.002 | 0.42 | 0.05  | 0.173 | -0.33  | -0.037 | -0.067 | 0.105 | -0.004 | -0.005 |

g

| Trend Removed | wAsia-JJA-1979-2013 |       |       |        |       |        |       |       |       |        |       |       |
|---------------|---------------------|-------|-------|--------|-------|--------|-------|-------|-------|--------|-------|-------|
|               | A                   |       |       | C-A    |       |        | A-C   |       |       | C      |       |       |
|               | OCC                 | PER   | DUR   | OCC    | PER   | DUR    | OCC   | PER   | DUR   | OCC    | PER   | DUR   |
| None          | 0.546               | 0.061 | 0.21  | -0.366 | -0.04 | -0.135 | 0.057 | 0.005 | -0.02 | -0.237 | 0.031 | 0.035 |
| N.Hemisphere  | 0.352               | 0.034 | 0.127 | -0.301 | -0.03 | -0.126 | 0.114 | 0.004 | 0     | -0.165 | 0.036 | 0.046 |

h

| Trend Removed | eNA-SO-1979-2013 |       |       |        |        |        |       |       |        |       |        |       |
|---------------|------------------|-------|-------|--------|--------|--------|-------|-------|--------|-------|--------|-------|
|               | A                |       |       | A-C    |        |        | C-A   |       |        | C     |        |       |
|               | OCC              | PER   | DUR   | OCC    | PER    | DUR    | OCC   | PER   | DUR    | OCC   | PER    | DUR   |
| None          | 0.438            | 0.019 | 0.086 | -0.285 | -0.011 | -0.063 | 0.097 | 0.013 | -0.037 | -0.25 | -0.007 | 0.003 |
| N.Hemisphere  | 0.366            | 0.017 | 0.075 | -0.266 | -0.005 | -0.055 | 0.091 | 0.01  | -0.041 | -0.19 | -0.002 | 0.003 |

i

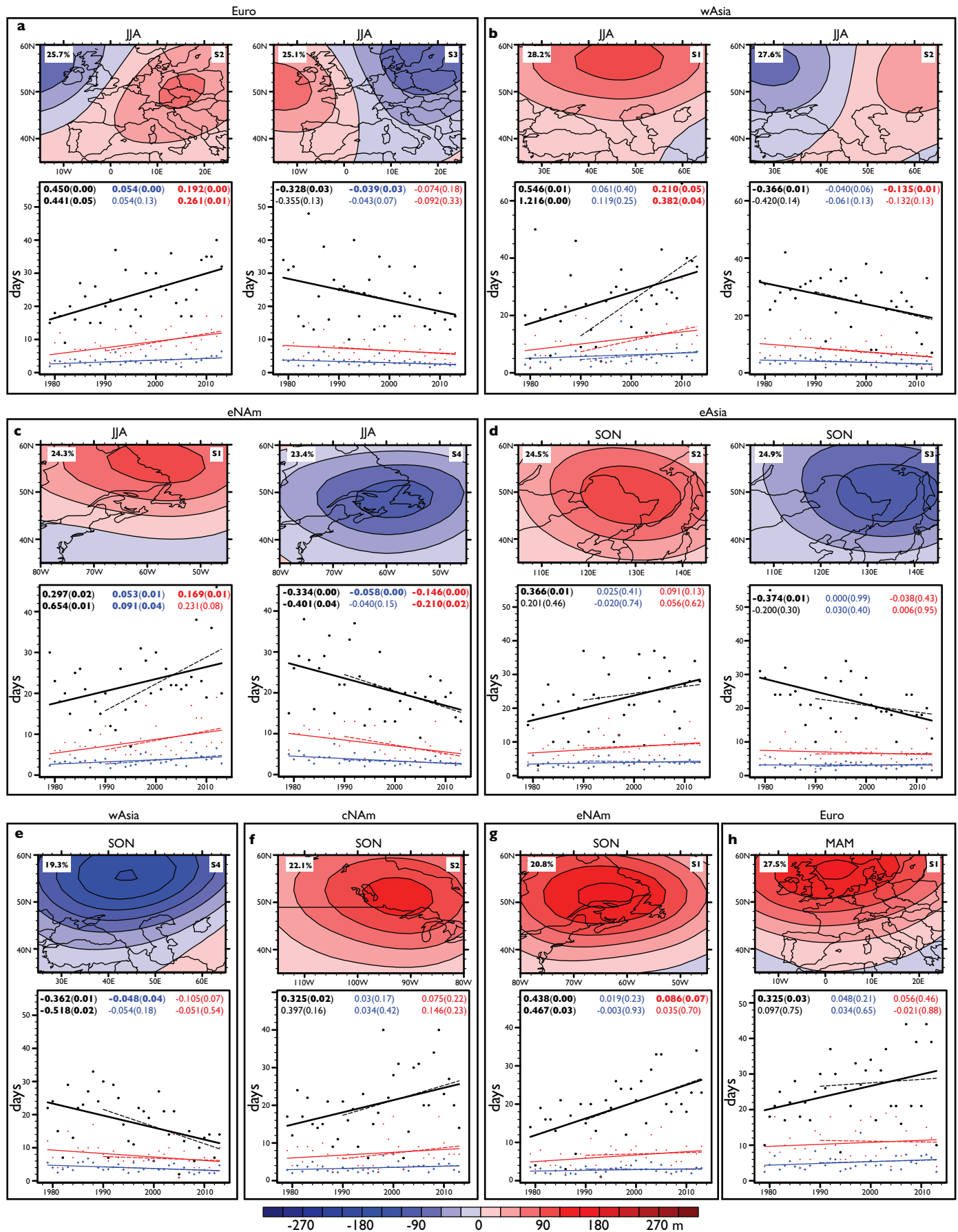
| Trend Removed | eAsia-SO-1979-2013 |        |       |       |       |       |        |       |        |       |       |       |
|---------------|--------------------|--------|-------|-------|-------|-------|--------|-------|--------|-------|-------|-------|
|               | A-C                |        |       | A     |       |       | C      |       |        | C-A   |       |       |
|               | OCC                | PER    | DUR   | OCC   | PER   | DUR   | OCC    | PER   | DUR    | OCC   | PER   | DUR   |
| None          | -0.114             | -0.01  | 0.015 | 0.366 | 0.025 | 0.091 | -0.374 | 0     | -0.038 | 0.122 | 0.02  | 0.049 |
| N.Hemisphere  | -0.101             | -0.003 | 0.015 | 0.217 | 0.01  | 0.036 | -0.229 | 0.034 | -0.017 | 0.113 | 0.023 | 0.049 |

j

| Trend Removed | cAsia-DJF-1990-2013 |       |       |       |       |       |        |       |        |        |        |        |
|---------------|---------------------|-------|-------|-------|-------|-------|--------|-------|--------|--------|--------|--------|
|               | A                   |       |       | A-C   |       |       | C-A    |       |        | C      |        |        |
|               | OCC                 | PER   | DUR   | OCC   | PER   | DUR   | OCC    | PER   | DUR    | OCC    | PER    | DUR    |
| None          | -0.127              | 0.03  | 0.127 | 1.044 | 0.119 | 0.262 | -0.621 | 0.013 | -0.07  | -0.296 | -0.005 | -0.042 |
| N.Hemisphere  | -0.169              | 0.037 | 0.103 | 1.105 | 0.144 | 0.319 | -0.701 | 0.021 | -0.056 | -0.235 | 0.004  | 0.006  |

**Extended Data Figure 3 | Geopotential height trends and thermal dilation adjustment.** **a–d**, Northern Hemisphere polar projections of 1979–2013 seasonal trends in 500 hPa geopotential heights (same as Fig. 1, reproduced here for convenience). **e**, Area-weighted trends in seasonal geopotential heights over the Northern Hemisphere and regional SOM domains. **f–j**, Trends in raw and detrended geopotential height SOM pattern occurrence (OCC), persistence (PER) and maximum duration (DUR) in

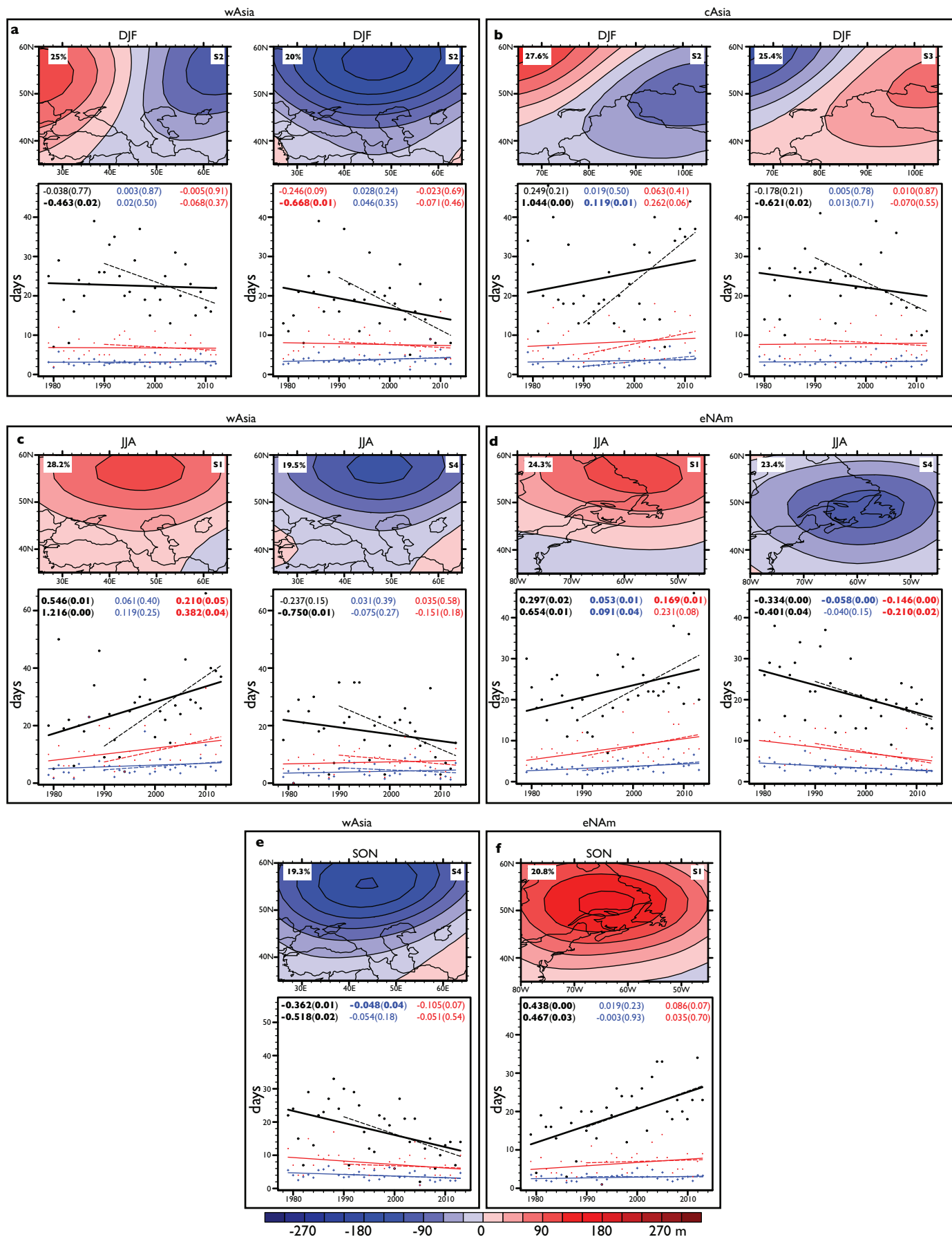
units of  $\text{d yr}^{-1} \text{yr}^{-1}$  for domains and seasons highlighted in the main text. The magnitudes of the (removed) seasonal Northern Hemisphere trends can be found in **e**. Grid cells highlighted in grey contain trends significant at the 5% level (Methods). SOM circulation patterns are abbreviated as follows: A, anticyclonic; C, cyclonic; and combinations of the two represent dipole patterns and west–east configurations. Geopotential height fields are sourced from the NCEP–DOE–R2 reanalysis<sup>33</sup>.





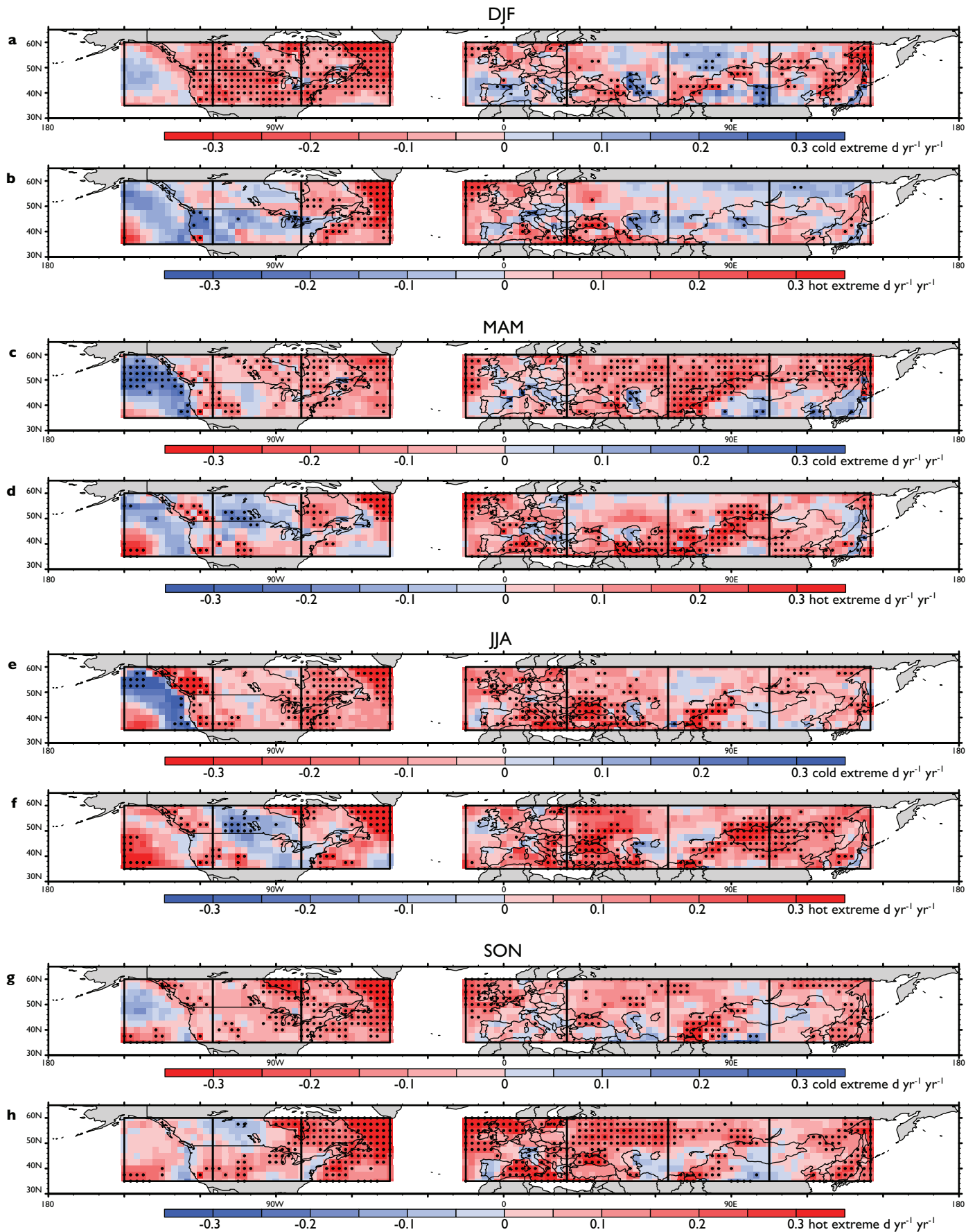
**Extended Data Figure 4 | 1979–2013 (satellite era) robust atmospheric circulation pattern trends.** Time series of circulation pattern occurrence (black ( $\text{d yr}^{-1}$ )), persistence (blue ( $\text{d event}^{-1}$ )) and maximum duration (red ( $\text{d event}^{-1}$ )) from the NCEP-DOE-R2 reanalysis<sup>33</sup>: **a**, summer over Europe; **b**, summer over western Asia; **c**, summer over eastern North America;

**d**, autumn over eastern Asia; **e**, autumn over western Asia; **f**, autumn over central North America; **g**, autumn over eastern North America; and **h**, spring over Europe. Statistically significant trends ( $(\text{yr}^{-1})$ ; 5% significance level; Methods) are identified by bold font in the scatter plots.



**Extended Data Figure 5 | 1990–2013 (ice era) robust atmospheric circulation pattern trends.** Time series of circulation pattern occurrence (black ( $\text{d yr}^{-1}$ )), persistence (blue ( $\text{d event}^{-1}$ )) and maximum duration (red ( $\text{d event}^{-1}$ )) from the NCEP-DOE-R2 reanalysis<sup>33</sup>: **a**, winter over western Asia;

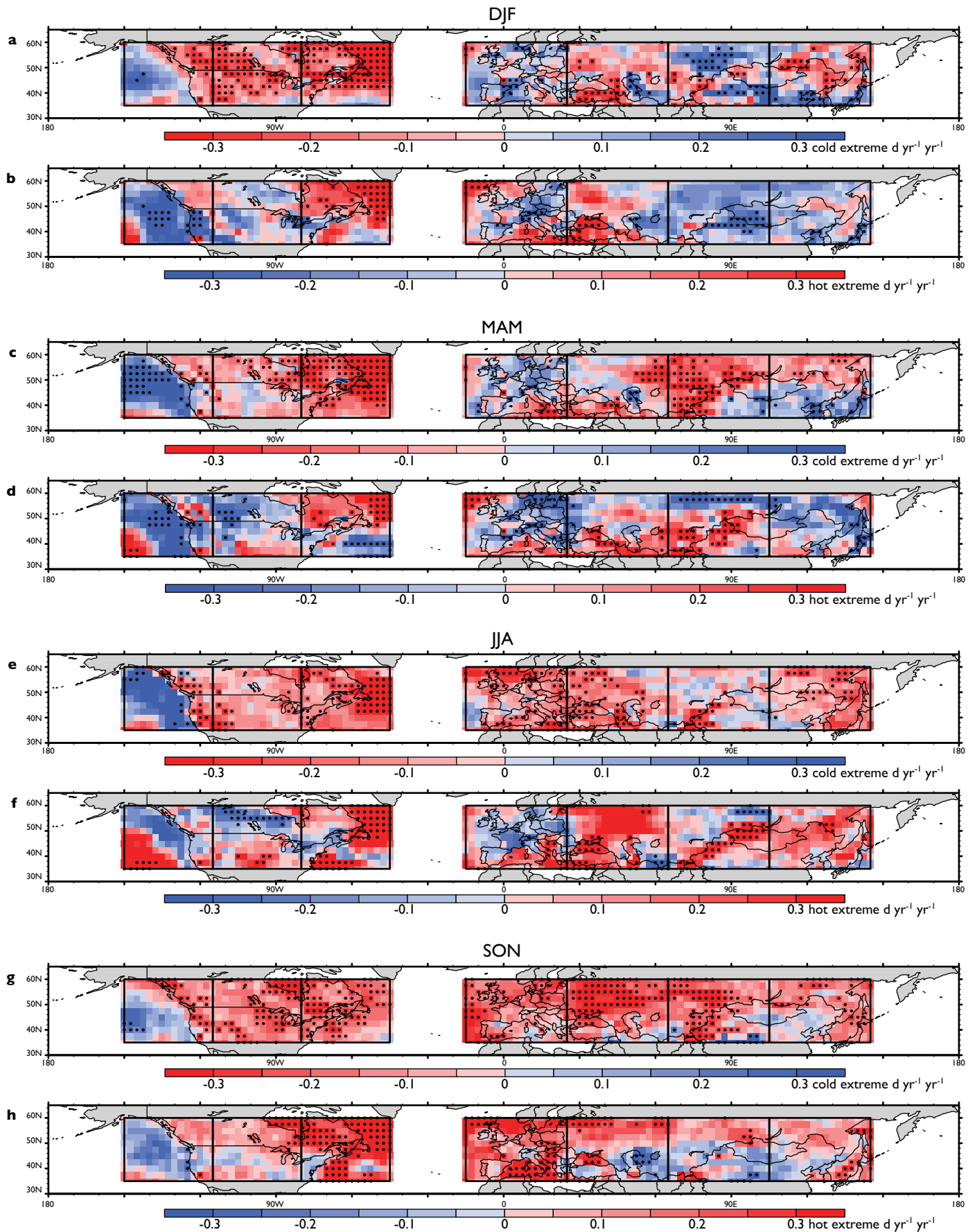
**b**, winter over central Asia; **c**, summer over western Asia; **d**, summer over eastern North America; **e**, autumn over western Asia; and **f**, autumn over eastern North America. Statistically significant trends ( $(\text{yr}^{-1})$ ; 5% significance level; Methods) are identified by bold font in the scatter plots.





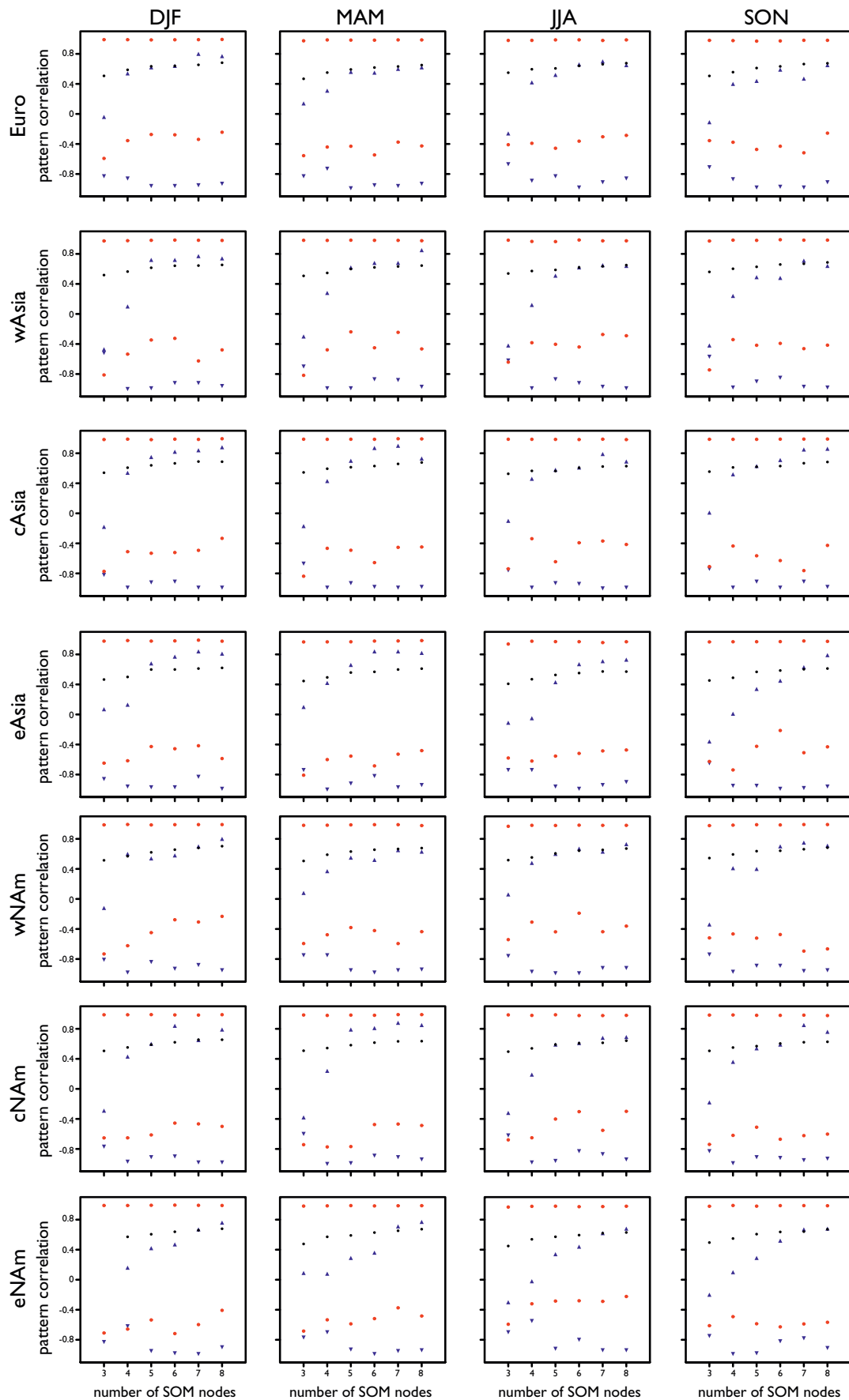
**Extended Data Figure 6 | 1979–2013 (satellite era) Northern Hemisphere extreme temperature occurrence trends.** Satellite-era extreme temperature trends ( $\text{d yr}^{-1} \text{yr}^{-1}$ ) for winter cold (**a**) and hot (**b**) occurrences; spring cold (**c**) and hot (**d**) occurrences; summer cold (**e**) and hot (**f**) occurrences; and

autumn cold (**g**) and hot (**h**) occurrences. Trends are calculated from the NCEP-DOE-R2 reanalysis 2-m daily maximum/minimum temperatures<sup>33</sup>. Grid cells with statistically significant trends (5% significance level; Methods) are stippled.



**Extended Data Figure 7 | 1990–2013 (ice era) Northern Hemisphere extreme temperature occurrence trends.** Ice-era extreme temperature trends ( $\text{d yr}^{-1} \text{ yr}^{-1}$ ) for winter cold (a) and hot (b) occurrences; spring cold (c) and hot (d) occurrences; summer cold (e) and hot (f) occurrences; and autumn cold

(g) and hot (h) occurrences. Trends are calculated from the NCEP-DOE-R2 reanalysis 2-m daily maximum/minimum temperatures<sup>33</sup>. Grid cells with statistically significant trends (5% significance level; Methods) are stippled.



● Max/Min pattern correlation of daily geopotential height anom. fields with matching SOM node pattern    ● Mean pattern correlation of all daily geopotential height anom. fields with matching SOM node patterns    ▲ Maximum SOM pattern to SOM pattern correlation    ▼ Minimum SOM pattern to SOM pattern correlation



**Extended Data Figure 8 | Sensitivity of pattern similarity to number of SOM nodes.** To determine an adequate number of SOM nodes, we follow a modified version of the methodology introduced by ref. 12, wherein the mean pattern correlation of all daily geopotential height anomaly fields and their matching SOM node patterns are computed for a suite of different SOM node counts (3, 4, 5, 6, 7 and 8), for all regions and all seasons (black dots). We also compute the maximum/minimum pattern correlation of daily geopotential height anomaly fields with their matching SOM node pattern (red dots) and the maximum/minimum SOM-pattern-to-SOM-pattern correlation (blue

triangles). The goal is to select an adequate number of nodes such that: (1) the mean pattern correlation of all daily geopotential height anomaly fields is relatively large; (2) the minimum pattern correlation of daily geopotential height anomaly fields is relatively large; and (3) the maximum SOM-pattern-to-SOM-pattern correlation is relatively small. Similar to ref. 12, we find that four SOM nodes are generally sufficient to capture the different modes of atmospheric variability, but small enough that SOM patterns depict distinct circulations. Geopotential height anomaly fields are sourced from the NCEP-DOE-R2 reanalysis<sup>33</sup>.

Extended Data Table 1 | Significant reanalysis circulation pattern trends and summary of multiple hypothesis testing

a

| Statistically Significant Trends ( $P<0.05$ ) |     |               |     |     |         |     |     |
|---|-----|---------------|-----|-----|---------|-----|-----|
|   | n   | satellite-era |     |     | ice-era |     |     |
|   |     | occ           | per | dur | occ     | per | dur |
| Linear Regression Analysis                    |     |               |     |     |         |     |     |
| NCAR/NCEP-R1                                  | 112 | 17            | 6   | 7   | 15      | 10  | 9   |
| NCEP-DOE-R2                                   | 112 | 16            | 5   | 7   | 13      | 8   | 9   |
| ERA-Interim                                   | 112 | 16            | 6   | 10  | 14      | 9   | 9   |
| Box-Cox Transformation Analysis               |     |               |     |     |         |     |     |
| NCAR/NCEP-R1                                  | 112 | 15            | 9   | 8   | 16      | 9   | 10  |
| NCEP-DOE-R2                                   | 112 | 15            | 6   | 7   | 16      | 10  | 10  |
| ERA-Interim                                   | 112 | 16            | 8   | 12  | 14      | 9   | 11  |

b

| Anderson-Darling Test for Normality                            |     |               |     |     |         |     |     |
|--|-----|---------------|-----|-----|---------|-----|-----|
|  | n   | satellite-era |     |     | ice-era |     |     |
|  |     | occ           | per | dur | occ     | per | dur |
| Original Data  |     |               |     |     |         |     |     |
| NCAR/NCEP-R1   | 112 | 103           | 47  | 56  | 102     | 76  | 68  |
| NCEP-DOE-R2  | 112 | 102           | 57  | 58  | 102     | 74  | 70  |
| ERA-Interim  | 112 | 99            | 62  | 49  | 105     | 76  | 65  |
| Original Data with FWER Multiple Hypothesis Testing            |     |               |     |     |         |     |     |
| NCAR/NCEP-R1   | 112 | 112           | 102 | 102 | 112     | 103 | 107 |
| NCEP-DOE-R2  | 112 | 112           | 100 | 103 | 112     | 107 | 109 |
| ERA-Interim  | 112 | 112           | 100 | 102 | 112     | 104 | 106 |
| Box-Cox Transformed Data                                       |     |               |     |     |         |     |     |
| NCAR/NCEP-R1   | 112 | 110           | 109 | 110 | 109     | 110 | 112 |
| NCEP-DOE-R2  | 112 | 111           | 110 | 110 | 110     | 111 | 112 |
| ERA-Interim  | 112 | 110           | 111 | 109 | 111     | 109 | 108 |
| Box-Cox Transformed Data with FWER Multiple Hypothesis Testing |     |               |     |     |         |     |     |
| NCAR/NCEP-R1   | 112 | 112           | 112 | 112 | 112     | 112 | 112 |
| NCEP-DOE-R2  | 112 | 112           | 112 | 112 | 112     | 112 | 112 |
| ERA-Interim  | 112 | 112           | 112 | 112 | 112     | 112 | 112 |

c

| Multiple Hypothesis Testing |       |               |     |     |         |     |     |
|-----------------------------|-------|---------------|-----|-----|---------|-----|-----|
| Control Type                | Level | satellite-era |     |     | ice-era |     |     |
|                             |       | occ           | per | dur | occ     | per | dur |
| NCAR/NCEP-R1                |       |               |     |     |         |     |     |
| FWER                        | 5%    | 2             | 0   | 0   | 0       | 1   | 0   |
|                             | 10%   | 2             | 0   | 1   | 1       | 1   | 0   |
| k-FWER                      | 5%    | 4             | 1   | 2   | 1       | 2   | 0   |
|                             | 10%   | 7             | 2   | 3   | 4       | 2   | 2   |
| FDR                         | 5%    | 0             | 0   | 0   | 0       | 0   | 0   |
|                             | 10%   | 2             | 0   | 0   | 0       | 0   | 0   |
| NCEP-DOE-R2                 |       |               |     |     |         |     |     |
| FWER                        | 5%    | 2             | 0   | 0   | 0       | 1   | 0   |
|                             | 10%   | 2             | 0   | 1   | 0       | 1   | 0   |
| k-FWER                      | 5%    | 2             | 1   | 2   | 2       | 2   | 0   |
|                             | 10%   | 5             | 2   | 4   | 5       | 3   | 0   |
| FDR                         | 5%    | 0             | 0   | 0   | 0       | 0   | 0   |
|                             | 10%   | 2             | 0   | 0   | 0       | 0   | 0   |
| ERA-Interim                 |       |               |     |     |         |     |     |
| FWER                        | 5%    | 0             | 0   | 1   | 0       | 0   | 0   |
|                             | 10%   | 1             | 0   | 1   | 0       | 0   | 0   |
| k-FWER                      | 5%    | 3             | 2   | 2   | 2       | 0   | 1   |
|                             | 10%   | 7             | 2   | 2   | 5       | 3   | 3   |
| FDR                         | 5%    | 0             | 0   | 0   | 0       | 0   | 0   |
|                             | 10%   | 0             | 0   | 0   | 0       | 0   | 0   |

a. We indicate the number of linear regression and Box-Cox transformed statistically significant occurrence (occ), persistence (per) and maximum duration (max dur) trends that surpass the 95% confidence threshold (5% significance level; Methods) for each reanalysis in each analysis period (1979–2013 is 'satellite-era' and 1990–2013 is 'ice-era'). For a given metric (occ, per and max dur) in a given period (satellite-era and ice-era), there are four patterns per region, over four seasons, for 7 domains. There are therefore 112 total trends in each metric in each period. b. Results of Anderson-Darling<sup>38</sup> tests of normality indicating the number of original and Box-Cox transformed<sup>40</sup> distributions of residuals that do not reject the null hypothesis of Gaussianity at the 5% level. For both the original and Box-Cox Anderson-Darling tests, we also apply multiple hypothesis testing to control the familywise error rate<sup>39</sup> (FWER) at the 5% significance level. c. To protect against the possibility of false positives in the linear regression trend analyses, for each reanalysis, we use three different multiple hypothesis error control methodologies to assess the number of locally significant hypotheses that surpass the global 5% or 10% significance level under strict error control requirements. FWER<sup>39</sup>, k-familywise error rate<sup>41</sup> (k-FWER), and false discovery rate<sup>39</sup> (FDR) analyses are applied (Methods). We note that all three metrics heavily favour the null. The presence of local tests that reject the null therefore represents a strong confirmation of the significance of those local tests, as those local tests are so highly significant that they can withstand the stricter multiple testing error control criteria.

Extended Data Table 2 | Quantitative partitioning of temperature extreme trends for select SOM analyses

|          |  |              |                       |                      |           |              |             |             |                  |                 |                |               |        |
|----------|--|--------------|-----------------------|----------------------|-----------|--------------|-------------|-------------|------------------|-----------------|----------------|---------------|--------|
| a        | 1979-2013 Europe JJA Hot Extreme Partitioning                |              |                       |                      |           |              |             |             |                  |                 |                |               |        |
|          | Pat occ trend  | Mean pat occ | T95 per pat occ trend | Mean T95 per pat occ | Dyn trend | Thermo trend | Inter trend | Total trend | Percent of trend | Percent dynamic | Percent thermo | Percent inter |        |
|          | SOM1   | -0.16        | 24.94                 | 0.001                | 0.07      | -0.01        | 0.02        | 0.00        | 0.01             | 10.1            | -112.0         | 237.9         | -25.9  |
|          | SOM2   | 0.45         | 23.66                 | 0.002                | 0.06      | 0.03         | 0.04        | 0.00        | 0.06             | 61.5            | 44.3           | 57.3          | -1.6   |
|          | SOM3   | -0.33        | 23.09                 | 0.001                | 0.03      | -0.01        | 0.02        | 0.00        | 0.01             | 8.7             | -109.3         | 200.9         | 8.4    |
|          | SOM4   | 0.04         | 20.31                 | 0.001                | 0.03      | 0.00         | 0.02        | 0.00        | 0.02             | 19.7            | 5.9            | 96.6          | -2.5   |
| Overall: |  |              |                       | 0.01                 | 0.10      | 0.00         | 0.10        |             |                  |                 |                |               |        |
| b        | 1979-2013 Western Asia JJA Hot Extreme Partitioning          |              |                       |                      |           |              |             |             |                  |                 |                |               |        |
|          | Pat occ trend  | Mean pat occ | T95 per pat occ trend | Mean T95 per pat occ | Dyn trend | Thermo trend | Inter trend | Total trend | Percent of trend | Percent dynamic | Percent thermo | Percent inter |        |
|          | SOM1   | 0.55         | 25.94                 | 0.002                | 0.07      | 0.04         | 0.04        | 0.03        | 0.11             | 70.3            | 36.5           | 37.8          | 25.7   |
|          | SOM2   | -0.37        | 25.40                 | 0.001                | 0.05      | -0.02        | 0.04        | 0.00        | 0.02             | 11.8            | -92.3          | 198.0         | -5.7   |
|          | SOM3   | 0.06         | 22.69                 | 0.001                | 0.03      | 0.00         | 0.02        | -0.01       | 0.02             | 12.0            | 9.9            | 122.3         | -32.1  |
|          | SOM4   | -0.24        | 17.97                 | 0.001                | 0.03      | -0.01        | 0.02        | -0.01       | 0.01             | 5.9             | -72.2          | 240.8         | -68.6  |
| Overall: |  |              |                       | 0.02                 | 0.12      | 0.01         | 0.16        |             |                  |                 |                |               |        |
| c        | 1979-2013 Eastern North America JJA Hot Extreme Partitioning |              |                       |                      |           |              |             |             |                  |                 |                |               |        |
|          | Pat occ trend  | Mean pat occ | T95 per pat occ trend | Mean T95 per pat occ | Dyn trend | Thermo trend | Inter trend | Total trend | Percent of trend | Percent dynamic | Percent thermo | Percent inter |        |
|          | SOM1   | 0.30         | 22.31                 | 0.002                | 0.07      | 0.02         | 0.04        | 0.01        | 0.07             | 51.4            | 29.9           | 56.8          | 13.3   |
|          | SOM2   | -0.02        | 21.63                 | 0.001                | 0.05      | 0.00         | 0.02        | 0.00        | 0.02             | 13.1            | -5.1           | 114.4         | -9.3   |
|          | SOM3   | 0.05         | 26.51                 | 0.001                | 0.04      | 0.00         | 0.04        | 0.00        | 0.03             | 26.7            | 6.8            | 106.7         | -13.6  |
|          | SOM4   | -0.33        | 21.54                 | 0.001                | 0.04      | -0.01        | 0.03        | 0.00        | 0.01             | 8.7             | -105.6         | 234.8         | -29.1  |
| Overall: |  |              |                       | 0.01                 | 0.12      | 0.00         | 0.13        |             |                  |                 |                |               |        |
| d        | 1990-2013 Central Asia DJF Cold Extreme Partitioning         |              |                       |                      |           |              |             |             |                  |                 |                |               |        |
|          | Pat occ trend  | Mean pat occ | T05 per pat occ trend | Mean T05 per pat occ | Dyn trend | Thermo trend | Inter trend | Total trend | Percent of trend | Percent dynamic | Percent thermo | Percent inter |        |
|          | SOM1   | -0.13        | 21.04                 | 0.001                | 0.04      | 0.00         | 0.01        | 0.00        | 0.01             | 9.0             | -79.0          | 177.8         | 1.3    |
|          | SOM2   | 1.04         | 24.61                 | 0.001                | 0.07      | 0.08         | 0.02        | 0.01        | 0.10             | 149.6           | 74.6           | 18.1          | 7.2    |
|          | SOM3   | -0.62        | 22.87                 | 0.000                | 0.02      | -0.01        | 0.01        | -0.01       | -0.01            | -21.3           | 98.0           | -45.4         | 47.4   |
|          | SOM4   | -0.30        | 21.74                 | -0.001               | 0.05      | -0.02        | -0.02       | 0.01        | -0.03            | -37.2           | 59.7           | 68.2          | -27.8  |
| Overall: |  |              |                       | 0.04                 | 0.02      | 0.01         | 0.07        |             |                  |                 |                |               |        |
| e        | 1979-2013 Eastern North America SON Hot Extreme Partitioning |              |                       |                      |           |              |             |             |                  |                 |                |               |        |
|          | Pat occ trend  | Mean pat occ | T95 per pat occ trend | Mean T95 per pat occ | Dyn trend | Thermo trend | Inter trend | Total trend | Percent of trend | Percent dynamic | Percent thermo | Percent inter |        |
|          | SOM1   | 0.44         | 18.89                 | 0.002                | 0.08      | 0.04         | 0.03        | 0.00        | 0.07             | 37.0            | 51.7           | 42.5          | 5.8    |
|          | SOM2   | -0.29        | 28.40                 | 0.001                | 0.03      | -0.01        | 0.03        | 0.00        | 0.02             | 10.5            | -45.3          | 156.9         | -11.6  |
|          | SOM3   | 0.10         | 25.91                 | 0.003                | 0.06      | 0.01         | 0.07        | 0.01        | 0.08             | 43.8            | 6.6            | 86.3          | 7.1    |
|          | SOM4   | -0.25        | 17.80                 | 0.002                | 0.03      | -0.01        | 0.03        | 0.00        | 0.02             | 8.7             | -44.1          | 161.9         | -17.8  |
| Overall: |  |              |                       | 0.03                 | 0.16      | 0.00         | 0.19        |             |                  |                 |                |               |        |
| f        | 1979-2013 East Asia SON Cold Extreme Partitioning            |              |                       |                      |           |              |             |             |                  |                 |                |               |        |
|          | Pat occ trend  | Mean pat occ | T05 per pat occ trend | Mean T05 per pat occ | Dyn trend | Thermo trend | Inter trend | Total trend | Percent of trend | Percent dynamic | Percent thermo | Percent inter |        |
|          | SOM1   | -0.11        | 19.43                 | -0.001               | 0.05      | -0.01        | -0.02       | 0.01        | -0.02            | 18.4            | 40.2           | 152.7         | -92.9  |
|          | SOM2   | 0.37         | 22.29                 | -0.001               | 0.03      | 0.01         | -0.01       | 0.00        | 0.00             | 0.8             | -1637.5        | 1844.1        | -106.5 |
|          | SOM3   | -0.37        | 22.69                 | -0.001               | 0.08      | -0.03        | -0.02       | 0.00        | -0.05            | 62.5            | 56.5           | 35.0          | 8.4    |
|          | SOM4   | 0.12         | 26.60                 | -0.001               | 0.04      | 0.00         | -0.02       | 0.00        | -0.02            | 18.3            | -31.2          | 119.7         | 11.5   |
| Overall: |  |              |                       | -0.02                | -0.07     | 0.01         | -0.08       |             |                  |                 |                |               |        |

Trends are partitioned for Europe (a), western Asia (b), and eastern North America (c) satellite-era summer hot extremes; central Asia ice-era winter cold extremes (d); eastern North America satellite-era autumn hot extremes (e); and eastern Asia satellite-era autumn cold extremes (f). From the left column moving right: (1) SOM node number; (2) trend in pattern occurrence ('pat-occ' d yr<sup>-1</sup>); (3) mean pattern occurrence for the period ('pat-occ' d yr<sup>-1</sup>); (4) trend in temperature extremes per pattern occurrence (intensity of extreme occurrence) ('ext-occ' d pat-occ d yr<sup>-1</sup>); (5) mean temperature extreme per pattern occurrence ('ext-occ' d pat-occ d yr<sup>-1</sup>); (6) trend in the dynamic influences ('ext-occ' d yr<sup>-1</sup> yr<sup>-1</sup>); (7) trend in the thermodynamic influences ('ext-occ' d yr<sup>-1</sup> yr<sup>-1</sup>); (8) trend in the interaction of dynamic and thermodynamic influences ('ext-occ' d yr<sup>-1</sup> yr<sup>-1</sup>); (9) the total trend in extreme occurrence for each SOM pattern ('ext-occ' d yr<sup>-1</sup> yr<sup>-1</sup>); (10) the per cent of the total trend in extreme occurrences that occur during each pattern (pattern trend percentages sum to 100%, meaning contributions from individual SOM patterns can be either positive or negative; that is, trend percentages greater than 100% indicate that other circulation patterns provide negative contributions); (11) the per cent of column 10 that is due to dynamic influences; (12) the per cent of column 10 that is due to thermodynamic influences; and (13) the per cent of column 10 that is due to interactive influences (dynamic, thermodynamic and interactive influence percentages for individual SOM patterns sum to 100%, meaning contributions from individual influences can be either positive or negative). The overall trends for the domain are presented below the individual SOM rows. All data are sourced from the NCEP-DOE-R2 reanalysis<sup>33</sup>.

# Sexual selection protects against extinction

Alyson J. Lumley<sup>1\*</sup>, Łukasz Michalczyk<sup>2\*</sup>, James J. N. Kitson<sup>1</sup>, Lewis G. Spurgin<sup>1</sup>, Catriona A. Morrison<sup>1</sup>, Joanne L. Godwin<sup>1</sup>, Matthew E. Dickinson<sup>1</sup>, Oliver Y. Martin<sup>3</sup>, Brent C. Emerson<sup>4</sup>, Tracey Chapman<sup>1</sup> & Matthew J. G. Gage<sup>1</sup>

**Reproduction through sex carries substantial costs, mainly because only half of sexual adults produce offspring<sup>1</sup>. It has been theorized that these costs could be countered if sex allows sexual selection to clear the universal fitness constraint of mutation load<sup>2–4</sup>. Under sexual selection, competition between (usually) males and mate choice by (usually) females create important intraspecific filters for reproductive success, so that only a subset of males gains paternity. If reproductive success under sexual selection is dependent on individual condition, which is contingent to mutation load, then sexually selected filtering through ‘genic capture’<sup>5</sup> could offset the costs of sex because it provides genetic benefits to populations. Here we test this theory experimentally by comparing whether populations with histories of strong versus weak sexual selection purge mutation load and resist extinction differently. After evolving replicate populations of the flour beetle *Tribolium castaneum* for 6 to 7 years under conditions that differed solely in the strengths of sexual selection, we revealed mutation load using inbreeding. Lineages from populations that had previously experienced strong sexual selection were resilient to extinction and maintained fitness under inbreeding, with some families continuing to survive after 20 generations of sib × sib mating. By contrast, lineages derived from populations that experienced weak or non-existent sexual selection showed rapid fitness declines under inbreeding, and all were extinct after generation 10. Multiple mutations across the genome with individually small effects can be difficult to clear, yet sum to a significant fitness load; our findings reveal that sexual selection reduces this load, improving population viability in the face of genetic stress.**

Sexual selection is a widespread evolutionary force giving rise to a striking diversity of sights, sounds and smells that filter reproductive success away from less competitive or attractive individuals, frequently at the expense of survival<sup>6</sup>. Sexual selection will operate to varying degrees whenever sexual reproduction exists, and its significance as a potent force profoundly influencing reproductive fitness of individuals is long established<sup>6</sup>. In contrast, limited empirical work has been directed at measuring the consequences of sexual selection for the fitness of populations. This lack of attention is surprising for two reasons: first, because population viability is vital for biodiversity maintenance and ecosystem stability, especially under modern anthropogenic stress<sup>7,8</sup>; and second, because it is predicted that the maintenance of costly sex as the dominant mode of reproduction might only be possible if it allows sexual selection to operate, reducing the universal handicap of mutation load<sup>3,4</sup>.

Population or lineage health will always suffer at some level from mutation load—the difference in fitness between a (usually theoretical) mutation-free lineage and one that carries a load of deleterious mutations that are segregating in mutation-selection balance<sup>7,9</sup>. This load exists because new mutations continually arise in all populations every generation, most of which will be deleterious<sup>10</sup>. Haldane calculated that mutation load would be unexpectedly high<sup>11</sup>, owing to large numbers of loci across each genome presenting multiple targets to unavoidable

mutation rates. Empirical estimates of mutation rate and load are hotly debated<sup>7</sup>, but we know, for example, that the average human lineage carries hundreds of deleterious loss-of-function mutations<sup>12</sup>, possibly thousands<sup>7</sup>. Natural selection will quickly remove mutations with large effects, but load persists through accumulation of mutations that have small individual effects, and/or exist as recessive alleles where their deleterious phenotypes are less frequently exposed to selection. Mutation load can therefore sum to a significant fitness constraint for a population which, because of its dispersed or concealed nature, is difficult to clear<sup>7,11</sup>. Sex could allow more effective purging of mutation load than asexual reproduction, if there are synergistic epistatic interactions between mutations<sup>7</sup>, so that their negative fitness impact is greater than strictly additive. However, the evidence that sufficient levels of synergistic epistasis exist remains equivocal<sup>7</sup>, so alternative explanations for the maintenance of costly sexual reproduction are sought.

Sexual selection could be a key filter against mutation load if, as Darwin acknowledged, “sexual selection will have given its aid to ordinary selection”<sup>13</sup>. This recognition that sexual selection places an additional, intraspecific filter on adaptive gene flow has been formalized by the idea of ‘genic capture’<sup>5</sup>, which proposes that reproductive success in the face of competition and choice depends on most or all aspects of an individual’s condition. Thus, sexual selection will act on most loci across the genome, purifying deleterious alleles from individuals within a lineage, and promoting fixation of advantageous ones, via three potentially connected routes. First, if competition or choice promotes non-random mating or fertilization success as an inverse consequence of mutation load, then deleterious dominant and recessive alleles will be under stronger purifying selection in lineages that experience sexual selection<sup>3</sup>. Second, if mutations are more deleterious in the competing sex than the choosing sex, which would occur if the competing sex suffers amplified reproductive success variance as a result of mutation load, then the equilibrium frequency of both dominant and recessive deleterious mutations will be lower within sexually selected populations, even accounting for synergistic epistasis in the competing sex<sup>2</sup>. Put simply, fathers should carry fewer mutations than all males<sup>3</sup>. Third, if female ability to maintain condition for reproductive fitness is under selection from interlocus sexual conflict, then sexual selection could purge mutations through females<sup>14,15</sup>. These routes for purging load via sexual selection provide theoretical explanations for how costly sex can persist as a dominant mode of reproduction<sup>2,3</sup>.

Few experiments have tested whether sexual selection removes deleterious mutations, yielding inconsistent findings. After introduction of mutations, heightened sexual selection rescued fitness faster in *Rhizoglyphus robini* mites<sup>16</sup>, *Drosophila melanogaster* flies<sup>17</sup> and *Onthophagus taurus* beetles<sup>18</sup>. However, further work showed that sexual selection reduced mutation loads only for a subset of fitness traits in *Drosophila*<sup>19</sup>. More detailed studies found no evidence that sexual selection could purge deleterious alleles from experimentally evolving lineages<sup>20</sup>, or restore fitness after mutation load had been induced, even after 60 generations<sup>21</sup>; in fact, reproductive fitness within

<sup>1</sup>School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. <sup>2</sup>Department of Entomology, Institute of Zoology, Jagiellonian University, Gronostajowa 9, 30-387 Kraków, Poland. <sup>3</sup>ETH Zurich, Institute of Integrative Biology, D-USYS, Universitätsstrasse 16, CHN J 11, 8092 Zürich, Switzerland. <sup>4</sup>Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), C/Astrofísico Francisco Sánchez 3, 38206 San Cristóbal de La Laguna, Santa Cruz de Tenerife, Canary Islands, Spain.

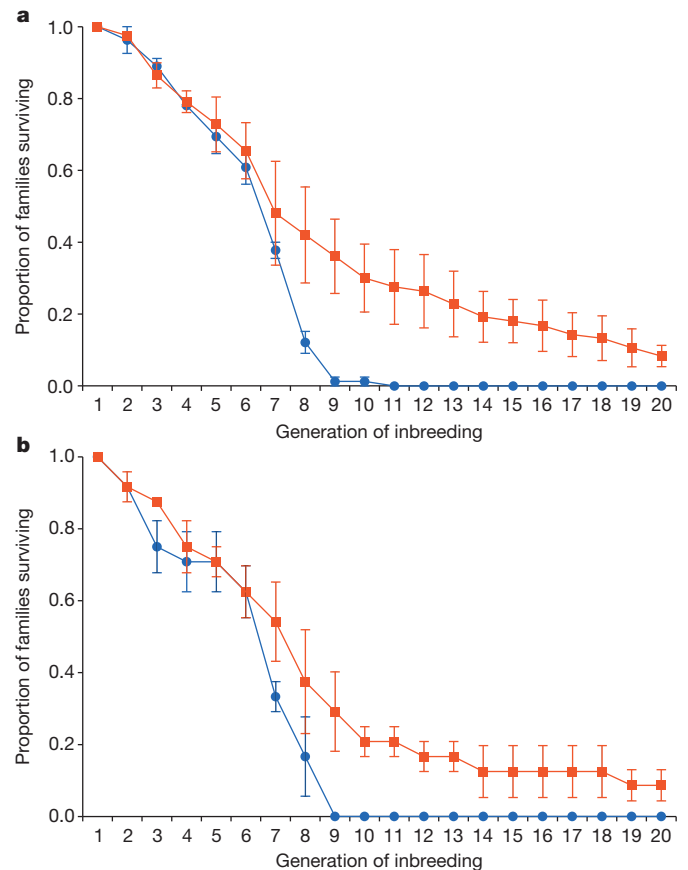
\*These authors contributed equally to this work.



mutated lines became higher when sexual selection had been minimized<sup>21</sup>. These inconsistent results could be explained by interlocus sexual conflict<sup>22</sup>: while sexual selection might play a beneficial role in purging mutation load, direct short-term constraints on population productivity may also arise when female fitness is constrained by conflicting adaptations that promote only male reproductive potential<sup>22</sup>. Thus, short-term fitness costs arising simultaneously from interlocus sexual conflict could confound the measurement of longer-term fitness benefits arising from sexual selection<sup>7</sup>. To avoid this problem, we assayed mutation load after populations (from the same ancestry) had experienced almost 7 years of sole variation in the intensity of sexual selection (Extended Data Figs 1 and 2). Although no evidence for carry-over effects of interlocus sexual conflict exists in our model under these conditions<sup>15</sup>, we also removed potential trans-generational effects by enforcing two generations of monogamous reproduction before beginning to assay fitness. To expose mutation load in each population, we enforced inbreeding using sib  $\times$  sib pairings, tracking fitness changes down multiple family lines (Extended Data Fig. 3). Because mutations will more likely persist as recessive alleles that are less frequently exposed to selection, compared with dominant wild types, fitness depression is a normal consequence of inbreeding as homozygosity increases the expression frequency of deleterious recessives through partial dominance<sup>23</sup>. By tracking extinction rates and fitness declines down inbreeding lineages perpetuated by monogamous sibling pairings, we tested the hypothesis that sex allows sexual selection to generate significant benefits through the purging of mutation load<sup>2,3</sup>, while avoiding the concurrent confound of interlocus sexual conflict<sup>22</sup>.

Experimental evolution lines began in 2005, with two different regimes (A and B) that both exposed replicate lineages to either 'strong' or 'weak' treatments of sexual selection, providing parallel independent experiments to measure whether sexual selection purges mutation load. Regime A exposed populations to 54 generations of divergent adult operational sex ratios that were either male-biased (10 females to 90 males) versus female-biased (90 females to 10 males), while regime B allowed polyandry (1 female to 5 males) versus enforced monogamy (1 female to 1 male) for 45 generations. All other conditions among lines and treatments within a regime were kept identical, including equalizing theoretical effective population sizes (see Methods, Extended Data Figs 2 and 4 and Supplementary Information for details confirming that heterozygosity was identical between treatments within either regime). The only difference between 'strong' and 'weak' treatments was therefore intensity of sexual selection during adult reproduction.

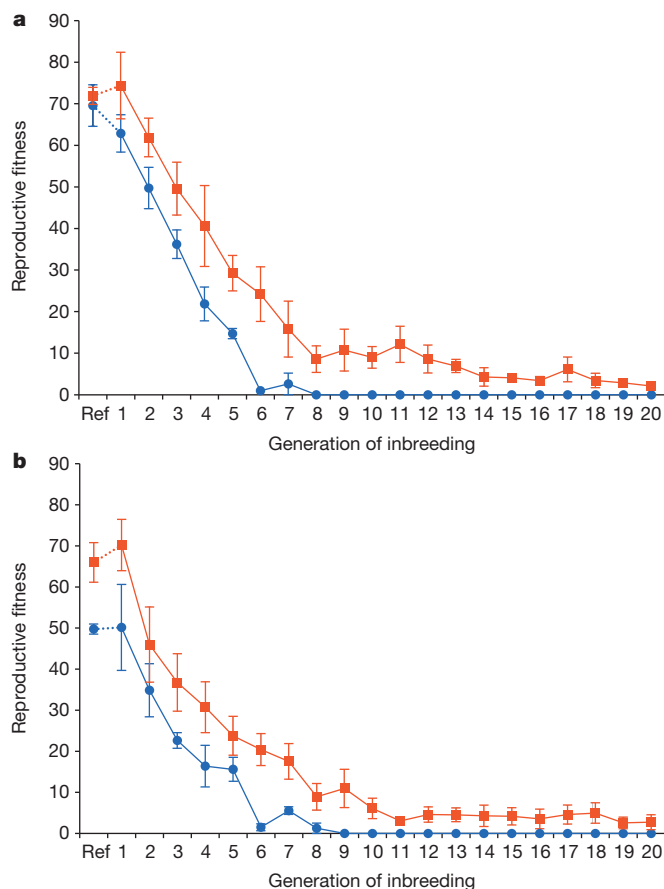
Our findings clearly showed for both regimes that lineages derived from populations experiencing evolutionary histories of strong sexual selection resisted extinction and maintained fitness more effectively when mutation load was exposed (Figs 1 and 2 and Extended Data Figs 5–7; all means presented in this article are arithmetic means  $\pm$  standard errors). Over the 3-year extinction assay, families derived from male-biased populations survived 44% longer than families from female-biased populations (mean number of generations to extinction =  $9.24 (\pm 1.29 \text{ s.e.m.})$  versus  $6.46 (\pm 0.15)$ ; Fig. 1a). Families from polyandrous histories survived 37% longer on average than those from monogamous treatments ( $8.50 (\pm 1.02)$  generations versus  $6.21 (\pm 0.46)$ ; Fig. 1b). When we combined data from regimes A and B into one analysis (incorporating sexual selection as a fixed factor variable), the history of sexual selection remained a significant predictor of number of generations to extinction ( $z = -3.43$ ,  $P < 0.001$ ), and there was no difference in extinction rates between regimes A and B ( $z = -0.51$ ,  $P = 0.611$ ), nor a significant interaction ( $z = 0.44$ ,  $P = 0.660$ ), revealing a consistent effect of sexual selection on extinction rates (Extended Data Fig. 5c). Overall, families derived from populations evolving under histories of strong sexual selection survived for 40% longer than those derived from weak sexual selection histories, giving an average survival time under mutation load exposure of  $8.87$



**Figure 1 | Extinction trajectories under increasing inbreeding differ between family lines derived from strong (red squares) versus weak (blue circles) sexual selection histories.** Each generation mean presents the average ( $\pm$  s.e.m.) proportion of surviving families for three independent lines per treatment. **a**, Regime A: male-biased (red) versus female-biased (blue) sexual selection histories. Each line is represented by 28 initial families ( $n = 84$  total families for either treatment). **b**, Regime B: polyandrous (red) versus monogamous (blue) selection histories. Each line is represented by eight initial families ( $n = 24$  total families for either treatment). Using parametric accelerated failure-time survival models with sexual selection treatment as a fixed effect, and incorporating correlated data within lines using a generalized estimating equation approach, we identified significantly lower extinction rates in populations that had previously experienced strong histories of sexual selection (Extended Data Fig. 5a, regime A:  $z = 3.40$ ,  $P < 0.001$ ; Extended Data Fig. 5b, regime B:  $z = 2.81$ ,  $P = 0.005$ ).

( $\pm 0.37$ ) versus  $6.33 (\pm 0.13)$  generations respectively. All 108 initial families derived from weak sexual selection histories ceased to produce offspring beyond the 10th generation of inbreeding, whereas 8 of the 108 families from the strong sexual selection histories were still producing offspring after 20 generations of inbreeding.

Declines in reproductive fitness exhibited similar patterns to the extinction rates they underpinned. Having removed sexual selection history via enforced monogamy in regime B, baseline reproductive fitness even without inbreeding is substantially reduced (Fig. 2b shows that a polyandrous history improves baseline fitness by  $\sim 30\%$ ). Most importantly, significant interactions between inbreeding and fitness decline revealed that fitness declines were much faster for families derived from treatments experiencing weak sexual selection histories in both regimes ( $z = 13.82$  and  $10.56$  for A and B respectively,  $P < 0.001$  for both; Fig. 2, Extended Data Table 1 and Extended Data Figs 6 and 7). To control for the possibility that failure-to-mate could have created differences between treatments, we repeated these analyses including only pairs that produced offspring: identical patterns of significant interactions between sexual selection history,



**Figure 2 | Reproductive fitness declines under increasing inbreeding of families derived from strong (red squares) versus weak (blue circles) sexual selection histories differ in magnitude and rate. a, Regime A: male-biased (red) versus female-biased (blue) selection histories. b, Regime B: polyandrous (red) versus monogamous (blue) histories.** Each generation mean presents the average ( $\pm$  s.e.m.) number of offspring produced under standardized conditions for three independent lines per treatment. Lines are represented by eight initial families ( $n = 24$  total families per treatment), and two breeding pairs per family. Average fitness under identical conditions but without inbreeding ('Ref') plotted for reference. Using generalized linear mixed models, and accounting for overdispersion as well as nesting non-independent replicate families within lines as random effects, we found in both regimes that inbreeding was a highly significant predictor of fitness (regime A:  $z = -21.34$ ,  $P < 0.001$ ; regime B:  $z = -21.17$ ,  $P < 0.001$ ), and that sexual selection treatment history was also a significant predictor of overall differences in fitness (regime A:  $z = -3.75$ ,  $P < 0.001$ ; regime B:  $z = -1.97$ ,  $P = 0.048$ ; see Extended Data Table 1 and Extended Data Figs 6 and 7).

inbreeding and fitness decline remained ( $z = 5.73$  and  $6.36$  for A and B respectively,  $P < 0.001$  for both; Extended Data Figs 6c, d and 7c, d).

Results from this 10-year experiment provide compelling empirical support for the complementary models given in refs 2 and 3, which argued that costs of sex<sup>1</sup> could be offset by population genetic benefits derived from sexual selection. The most obvious mechanism explaining these differences is that heightened sexual selection, via 'genetic capture'<sup>4,5</sup>, more effectively strips out, or prevents, mutation load from becoming fixed in a population, strengthening its ability to withstand stress. Hemizygous selection could make sex-linked loci especially prone to this process<sup>24</sup>, but since the *T. castaneum* X chromosome makes up only ~6% of the genome<sup>25</sup>, we expect purging of load to have occurred on the autosomes too. There is emerging evidence that sexual selection can profoundly shape the genome, with feminized patterns of sex-linked expression in the transcriptomes of both sexes after experimental evolution under monogamy in *Drosophila*<sup>26</sup>. Our findings indicate that sexual selection also acts to purge mutation load,

even in populations that we expect to be adapting close to their natural fitness peaks<sup>27</sup>. Stronger sexual selection will drive greater variance in reproductive success, so that the average father should carry fewer deleterious mutations than the average male<sup>3</sup>, and perhaps also the average mother if interlocus sexual conflict constrains female fitness<sup>15</sup>.

Within the promiscuous *Tribolium* model<sup>28</sup> (Extended Data Fig. 8 and Methods), both pre- and post-copulatory processes for winning or controlling fertilizations will operate through broad behaviours and physiology within the whole organism, down to competition and choice at the gamete level. Genic capture predicts that individual success across this continuum will depend on overall condition and genes at many loci<sup>5</sup>. Under monogamy, where sexual selection has been removed, populations suffer the constraints of sexual reproduction, but none of the benefits of sexual selection, explaining the lowered baseline fitness even without inbreeding, as well as more rapid fitness declines and heightened extinction rates under inbreeding (Figs 1 and 2 and Extended Data Figs 5–7). Although we exposed mutation load for experimental measurement through inbreeding, the intense and diverse demands from selection across multiple generations in the natural environment, where populations are more likely to be displaced from their fitness peaks<sup>27</sup>, are likely to expose even greater fitness differentials due to load variance. After only one generation of inbreeding, for example, mean fitness between strong versus weak sexual selection histories differs by 20% and 40% in regimes A and B respectively (Fig. 2). Our results indicate, as demands rise upon populations that have been depleted and fragmented, and displaced further from their fitness peaks<sup>27</sup> via increasing anthropogenic stress<sup>29</sup>, that sexual selection could be an important force protecting species or populations from the extinction vortex<sup>30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 January; accepted 18 March 2015.

Published online 18 May; corrected online 24 June 2015 (see full-text HTML version for details).

- Maynard Smith, J. *The Evolution of Sex* (Cambridge Univ. Press, 1978).
- Agrawal, A. F. Sexual selection and the maintenance of sexual reproduction. *Nature* **411**, 692–695 (2001).
- Siller, S. Sexual selection and the maintenance of sex. *Nature* **411**, 689–692 (2001).
- Whitlock, M. C. & Agrawal, A. F. Purging the genome with sexual selection: reducing mutation load through selection on males. *Evolution* **63**, 569–582 (2009).
- Tomkins, J. L., Radwan, J., Kotiaho, J. S. & Tregenza, T. Genic capture and resolving the lek paradox. *Trends Ecol. Evol.* **19**, 323–328 (2004).
- Andersson, M. B. *Sexual Selection* (Princeton Univ. Press, 1994).
- Agrawal, A. F. & Whitlock, M. C. Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. *Annu. Rev. Ecol. Syst.* **43**, 115–135 (2012).
- Holman, L. & Kokko, H. The consequences of polyandry for population viability, extinction risk and conservation. *Phil. Trans. R. Soc. B* **368**, <http://dx.doi.org/10.1098/rstb.2012.0053> (2013).
- Crow, J. F. Mutation, mean fitness, and genetic load. *Oxf. Surv. Evol. Biol.* **9**, 3–42 (1993).
- Keightley, P. D. & Lynch, M. Toward a realistic model of mutations affecting fitness. *Evolution* **57**, 683–685 (2003).
- Haldane, J. B. S. The effect of variation of fitness. *Am. Nat.* **71**, 337–349 (1937).
- McVean, G. A. et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
- Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
- Wigby, S. & Chapman, T. Female resistance to male harm evolves in response to manipulation of sexual conflict. *Evolution* **58**, 1028–1037 (2004).
- Michalczyk, L. et al. Experimental evolution exposes female and male responses to sexual selection and conflict in *Tribolium castaneum*. *Evolution* **65**, 713–724 (2011).
- Radwan, J. Effectiveness of sexual selection in removing mutations induced with ionizing radiation. *Ecol. Lett.* **7**, 1149–1154 (2004).
- Hollis, B., Fierst, J. L. & Houle, D. Sexual selection accelerates the elimination of a deleterious mutant in *Drosophila melanogaster*. *Evolution* **63**, 324–333 (2009).
- Almbro, M. & Simmons, L. W. Sexual selection can remove an experimentally induced mutation load. *Evolution* **68**, 295–300 (2014).
- Sharp, N. P. & Agrawal, A. F. Mating density and the strength of sexual selection against deleterious alleles in *Drosophila melanogaster*. *Evolution* **62**, 857–867 (2008).

20. Arbuthnott, D. & Rundle, H. D. Sexual selection is ineffectual or inhibits the purging of deleterious mutations in *Drosophila melanogaster*. *Evolution* **66**, 2127–2137 (2012).
21. Hollis, B. & Houle, D. Populations with elevated mutation load do not benefit from the operation of sexual selection. *J. Evol. Biol.* **24**, 1918–1926 (2011).
22. Holland, B. & Rice, W. R. Experimental removal of sexual selection reverses intersexual antagonistic coevolution and removes a reproductive load. *Proc. Natl Acad. Sci. USA* **96**, 5083–5088 (1999).
23. Charlesworth, B. & Charlesworth, D. The genetic basis of inbreeding depression. *Genet. Res.* **74**, 329–340 (1999).
24. Rice, W. R. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**, 735–742 (1984).
25. Kim, H. S. *et al.* BeetleBase in 2010: revisions to provide comprehensive genomic information for *Tribolium castaneum*. *Nucleic Acids Res.* **38**, D437–D442 (2010).
26. Hollis, B., Houle, D., Yan, Z., Kawecki, T. J. & Keller, L. Evolution under monogamy feminizes gene expression in *Drosophila melanogaster*. *Nature Commun.* **5**, <http://dx.doi.org/10.1038/ncomms4482> (2014).
27. Long, A. F. T., Agrawal, A. F. & Rowe, L. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. *Curr. Biol.* **22**, 204–208 (2012).
28. Michalczyk, Ł. *et al.* Inbreeding promotes female promiscuity. *Science* **333**, 1739–1742 (2011).
29. Lande, R. Genetics and demography in biological conservation. *Science* **241**, 1455–1460 (1988).
30. Gilpin, M. E. & Soulé, M. E. in *Conservation Biology: The Science of Scarcity and Diversity* (ed. Soulé, M. E.) 19–34 (Sinauer Associates, 1986).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank the Natural Environment Research Council and the Leverhulme Trust for financial support, D. Edward for statistical advice and colleagues at the 2013 Biology of Sperm meeting for comments that improved analytical design and interpretation.

**Author Contributions** Ł.M., O.Y.M. and M.J.G.G. initiated the experimental evolution lines used in this work in 2005 and, with A.J.L., have maintained them since. M.J.G.G., Ł.M. and A.J.L. conceived, designed, conducted and analysed the study, with input from B.C.E. and T.C. J.J.N.K. and L.G.S. ran the microsatellite analyses. J.L.G., M.E.D. and O.Y.M. helped with line maintenance and experimental data collection. C.A.M. performed the fitness analyses. M.J.G.G. and A.J.L. wrote the paper, with contributions from all authors.

**Author Information** Data sets for all experiments and assays have been deposited in the Dryad Digital Repository at <http://dx.doi.org/10.5061/dryad.86750>. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.J.G.G. ([m.gage@uea.ac.uk](mailto:m.gage@uea.ac.uk)).

## METHODS

**Experimental evolution lines.** Beetles were of the widely used Georgia-1 (GA1) 'wild type' strain, originally collected from stored corn in 1980, and since cultured by the Beeman laboratory (United States Department of Agriculture), maintained under standard conditions. Adult virgin beetles from the same ancestral GA1 population were randomly allocated to begin their respective treatments within both regimes. Each generation, male and female pupae were separated and placed in fresh fodder (organic flour, yeast (10:1) and oats) for 10 days to allow adult emergence and sexual maturation<sup>15,28,31</sup>. Then, using the controlled sex ratios in the different sexual selection treatments, mature adults were placed in *ad libitum* fodder for 7 days to compete, choose, mate, oviposit and therefore reproduce under divergent intensities of sexual selection<sup>15,28,31</sup>. After 7 days, adults were removed and the eggs and larvae (typically about 70 offspring per female) were left to develop under standardized conditions with equal offspring densities by maintaining *ad libitum* levels of food in proportion to the number of offspring. These conditions were maintained for about 7 years, allowing the application of strong versus weak sexual selection in a total of 12 independent lines through sole variation in the adult operational sex ratio, while equalizing effective population size within regimes. Two parallel experimental evolution regimes were run, regimes A and B, which both applied treatments that created contrasting divergences in sexual selection intensities (see Extended Data Fig. 2 for visual details). Under regime A, the male-biased treatment provided 10 females with choice among 90 males, and simultaneous competition between the 90 males to fertilize the 10 females, whereas in female-biased treatments the reverse scenarios applied. Six independent lines (three per treatment) were maintained for 54 generations under regime A. Regime B generated divergence by enforcing either monogamy (20 replicate pairs per line), or allowing reproduction to be achieved through competition between five males under polyandry (12 replicate groups per line). Thus, polyandry provided each female with a choice among five males, and competition between five males for just one female, whereas monogamy completely removed all female choice and male-male competition. Six independent lines (three per treatment) were maintained for 45 generations under regime B. After each adult reproduction period and sexual selection treatment, offspring from replicate families within each independent line were pooled under standardized densities for larval development and genetic mixing to the next generation.

**Quantifying sexual selection intensity.** Divergence in the strength of sexual selection between the different treatments was estimated using a male mating potential assay, in which males from the ancestral GA1 control stock population were each provided with a series of 84 unmated females. Individual were sexed as pupae and matings took place after 10 days of adult maturation (as per the sexual selection treatments). Males were placed individually in microcentrifuge tubes with approximately 1 ml of flour topped with oats; 24 h before their trial they were identified with a small dot of white correction fluid on their thorax<sup>15,28,31</sup>. Females were placed in groups of six per Petri dish with 7 g of flour topped with oats. Mating assays were performed by placing individual males into Petri dishes containing standard fodder and six females, moving males to a new group of six females every 12 h for 7 days ( $12 \times 7 = 84$  females given to each male over the week). After each 12 h mating period, females were placed singly into Petri dishes with standard fodder for 7 days to lay eggs. Any offspring were then allowed to develop for 35 days before being frozen, and their presence noted to score successful matings. Eleven males were assayed, and the average number of successfully fertilized females across this 1 week period per male was  $50 (\pm 3 \text{ s.e.m.})$ ; Extended Data Fig. 8), providing extreme divergence in potential levels of female polyandry (and therefore sexual selection) between strong and weak sexual selection treatments. Thus, in regime A the 10 females at each adult generation in the male-biased treatment were potentially exposed to  $90 \times 50 = 4,500$  successful matings, or 450 matings per female per week involving all 90 males. Whereas the 90 females in the female-biased treatment potentially experienced 500 matings ( $10 \times 50$ ), or 5.6 per female per week, enabling each female to be mated, on average, by 5.6 of the males available. In the monogamous treatment, we removed sexual selection altogether as only one male was available for mating, while the polyandrous treatment generated a fivefold increase in potential for pre- and post-copulatory male-male competition and female choice through the availability of five males, all of which could mate with the female.

**Effective population sizes.** To avoid differential inbreeding that could have subsequently influenced our extinction assay, we equalized the theoretical effective population size,  $N_e$ , for mixed adult sex ratios ( $N_e = \frac{4N_f N_m}{N_f + N_m}$ ) (ref. 32) in our divergent sexual selection treatments; thus, in regime A  $N_e = 36$  for both male- and female-biased treatments, and in regime B  $N_e = 40$  for both monogamous and polyandrous treatments. It is important to note that strong sexual selection treatments with male-biased or polyandrous structures may translate into reduced realized  $N_e$  due to male success skew under strong sexual selection (where competition and choice could allow reproductive success by a smaller subset of males

in the breeding population). Thus, our experimental design is conservative to any influence of  $N_e$ , because we test the prediction that lineages with histories of strong sexual selection (and therefore potentially lower  $N_e$ ) should demonstrate resistance to extinction and reduced rates of fitness decline under inbreeding (because they have more effectively purged mutation load). If  $N_e$  is lower in these strong sexual selection populations, we would expect this to result in reduced heterozygosity at the start of inbreeding, resulting in lower initial fitness and therefore faster extinction rates (yet we observe the converse in Figs 1 and 2). As an additional, third, check for  $N_e$  confounds, we also directly measured heterozygosity and allelic richness in our lines to establish that differential inbreeding had not occurred under different sexual selection treatments (methods outlined below). By screening 628 individuals representing the different sexual selection treatments in both regimes A and B at 13 microsatellite loci (Supplementary Information), we were able to confirm that heterozygosity and allelic richness showed no differences between 'weak' or 'strong' sexual selection histories (Extended Data Fig. 4).

**Exposing mutation load.** Having been subjected to experimental evolution that applied strong versus weak sexual selection under equal effective population sizes, we then exposed the mutation load carried in these populations by inbreeding down multiple replicate family lines to expose deleterious recessives, and tracking fitness decline to extinction as a result of inbreeding depression through partial dominance<sup>23</sup>. No statistical methods were used to predetermine sample size, which was fundamentally controlled by the number of independent selection lines that had been maintained for 6 to 7 years. At the end of experimental evolution, we created multiple full-sib families using monogamous crosses of randomly chosen unmated males and females within each independent selection line, which simultaneously removed any transgenerational interlocus sexual conflict effects, and from which full-sib offspring were then used to continue a total of 216 family lines down increasing inbreeding coefficients and homozygosity (Extended Data Fig. 3). There were thus two generations of monogamy applied before fitness assays began, eliminating any carry-over effects of sexual selection or interlocus conflict. As levels of inbreeding increased, we tracked extinction rates and fitness declines for multiple families within each of the 12 independent lines from the original sexual selection treatments. In regime A, we created 28 full-sib families in each independent line ( $N_{\text{total}} = 28_{\text{families}} \times 3_{\text{independent lines}} \times 2_{\text{sexual selection treatments}} = 168$  families), allowing extinction rates to be measured within each of the three independent lines in either of the two sexual selection treatments. When sibling pairs failed to produce female and male offspring that survived to adulthood to enable continuation of the line, it was recorded as extinct (see Extended Data Fig. 3). In addition to extinction, we also measured reproductive fitness in a subset of these families, assayed as the average number of offspring produced by two randomly chosen sib  $\times$  sib pairs within eight of the families per line (Extended Data Fig. 3). Identical protocols for regime B were followed, except that both extinction rates and average fitness were measured across eight families per independent line ( $N_{\text{total}} = 8_{\text{families}} \times 3_{\text{independent lines}} \times 2_{\text{sexual selection treatments}} = 48$  families). Blinding was not performed as the protocol did not permit biasing to affect results. Inbreeding was continued for 20 generations of sib  $\times$  sib matings (which operated over 3 years), by which time Wright's inbreeding coefficient  $F$  had increased to 0.986. After 20 generations of inbreeding, 208 of the 216 initial lines had gone extinct, with the eight survivors all derived from strong sexual selection histories.

**Baseline fitness.** At four time-points throughout the extinction assay (parental,  $F_7$ ,  $F_{15}$  and  $F_{21}$  generations), we assayed 'baseline' fitness (without inbreeding) of standard monogamous crosses from the different sexual selection treatments, using identical protocols to fitness measures applied for the inbred crosses. Pupae from each of the three independent lines within each of the four different selection treatments were separated into single-sex groups and isolated for 10 days to allow adult eclosion and sexual maturation. Twenty randomly chosen male-female pairs were then established for each of the three independent lines within each of the four sexual selection treatments. Each pair was placed into a 7 ml plastic vial, with *ad libitum* food and oats, and allowed to mate and oviposit for 7 days. After this, the flour from each vial, containing the eggs laid during the 7 day mating period, was transferred to a Petri dish containing a further 10 g of fodder, and the eggs allowed to develop to adult eclosion ( $\sim 35$  days). Fitness was scored as the total number of adults produced from each cross across 7 days of mating and oviposition. Data from previous work<sup>15</sup> where male-female pairs were allowed to interact for 7 days shows (a) that the first week of oviposition produces  $>25\%$  of the total female reproductive fitness for such male-female interactions, and (b) that female fitness over the first week significantly correlates with total female reproductive fitness ( $r = 0.48$ ,  $P = 0.008$ ,  $n = 29$  pairs). Average baseline fitness values are presented for reference in Fig. 2a, b ('Ref').

**Extinction analyses.** Extinction rates were analysed using the 'survival' package<sup>33</sup> in R version 3.1.0 (ref. 34). To assess whether sexual selection history influenced extinction rate after 20 generations of sib  $\times$  sib inbreeding for either regimes A or B, we analysed generation to extinction using a parametric accelerated failure-time



survival model with Weibull baseline hazard distribution, taking into account right-censored data from families that were still alive at the end of the experiment. Generation of extinction was modelled with sexual selection treatment as a fixed effect, and the shape and scale of the underlying Weibull hazard distribution was allowed to vary by treatment. We also ran a combined analysis where both experimental evolution regimes were incorporated into the model, with level of sexual selection (strong versus weak), comparison (regime A or B) and their interaction as fixed effects. In all survival analyses, correlated data within lines were incorporated using a generalized estimating equation approach. Akaike's information criteria and graphical interpretation of the complementary log-log survival plots were used to confirm appropriateness of the specified Weibull distribution; no violations of model assumptions were detected. Model fits are presented in Extended Data Fig. 5a–c.

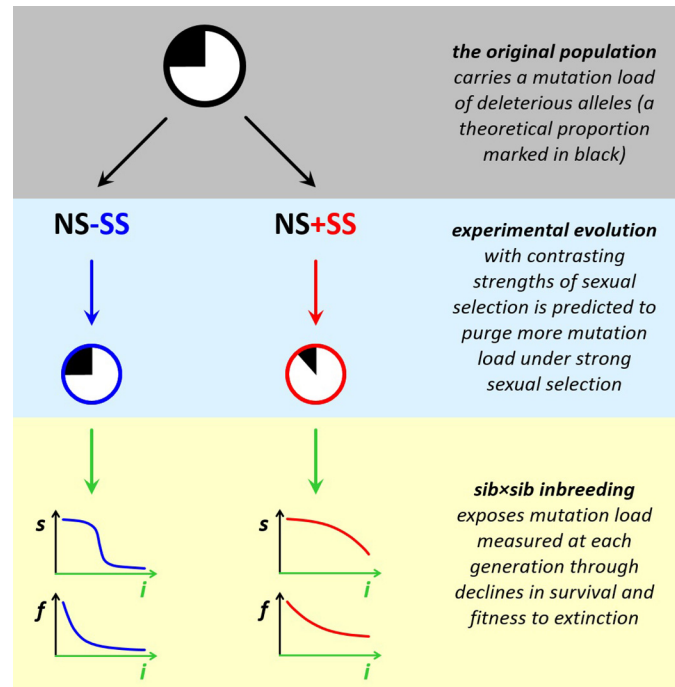
**Reproductive fitness analyses.** We tested for a relationship between fitness and generation of inbreeding by fitting generalized linear mixed models (GLMMs) using the glmmADMB package<sup>35</sup> in R 3.0.3 (ref. 34). We assumed a negative binomial error structure which added the parameter ( $k$ ) to the variance mean relationship, allowing us to account for the overdispersion in our data introduced by a large number of zero fitness observations. GlmmADMB has two options for fitting the relationship between the mean and the variance. These are family = 'nbinom1', which assumes the variance =  $k \times \text{mean}$ , and family = 'nbinom2', which assumes the variance =  $\text{mean}(1 + \text{mean}/k)$ . We fitted models with each of these options in turn and then selected the parameterization that provided the lowest AIC score and therefore accounted for the greatest amount of variance in our data set. To account for the nested nature of the experimental design and the non-independence between replicate families, random effects were included in the model as replicate nested within family nested within line. Separate GLMMs were fitted to the regime A (male-biased versus female-biased) comparison and the regime B (polyandry versus monogamy) comparison (Extended Data Figs 6 and 7). To ensure that failure-to-mate was not a reason for differences in fitness declines between treatments, we repeated the analyses including only those families that produced some offspring each generation (Extended Data Figs 6c, d and 7c, d).

**Microsatellite analyses.** A total of 628 individuals from the 12 independent lines representing all four sexual selection treatments (strong versus weak in regimes A and B) were genotyped at 13 loci conforming to Hardy–Weinberg equilibrium

(Supplementary Information). Scored genotypes for all individuals were analysed in Arlequin version 3.5 (ref. 36) to obtain allelic richness and observed and expected heterozygosity. We tested for differences in heterozygosity across selection treatments using linear mixed models, implemented in lme4 version 1.1-6 (ref. 37) package in R<sup>34</sup>, with significance testing performed using the package lmerTest version 2.0-6 (ref. 38). Heterozygosity was included as a response variable with selection treatment as a fixed factor and locus as a random factor. To assess whether independent line identity had a significant effect on genetic diversity beyond that of selection treatment, a second mixed model was tested with the same variables plus line identity as an additional random factor. The two models were then compared using likelihood ratio tests. All model residuals were tested for normality and no violations of model assumptions were found. Estimated heterozygosity from the linear mixed effect models and the associated standard error is plotted in Extended Data Fig. 4. Microsatellite data are available from the Dryad Digital Repository at <http://dx.doi.org/10.5061/dryad.86750>.

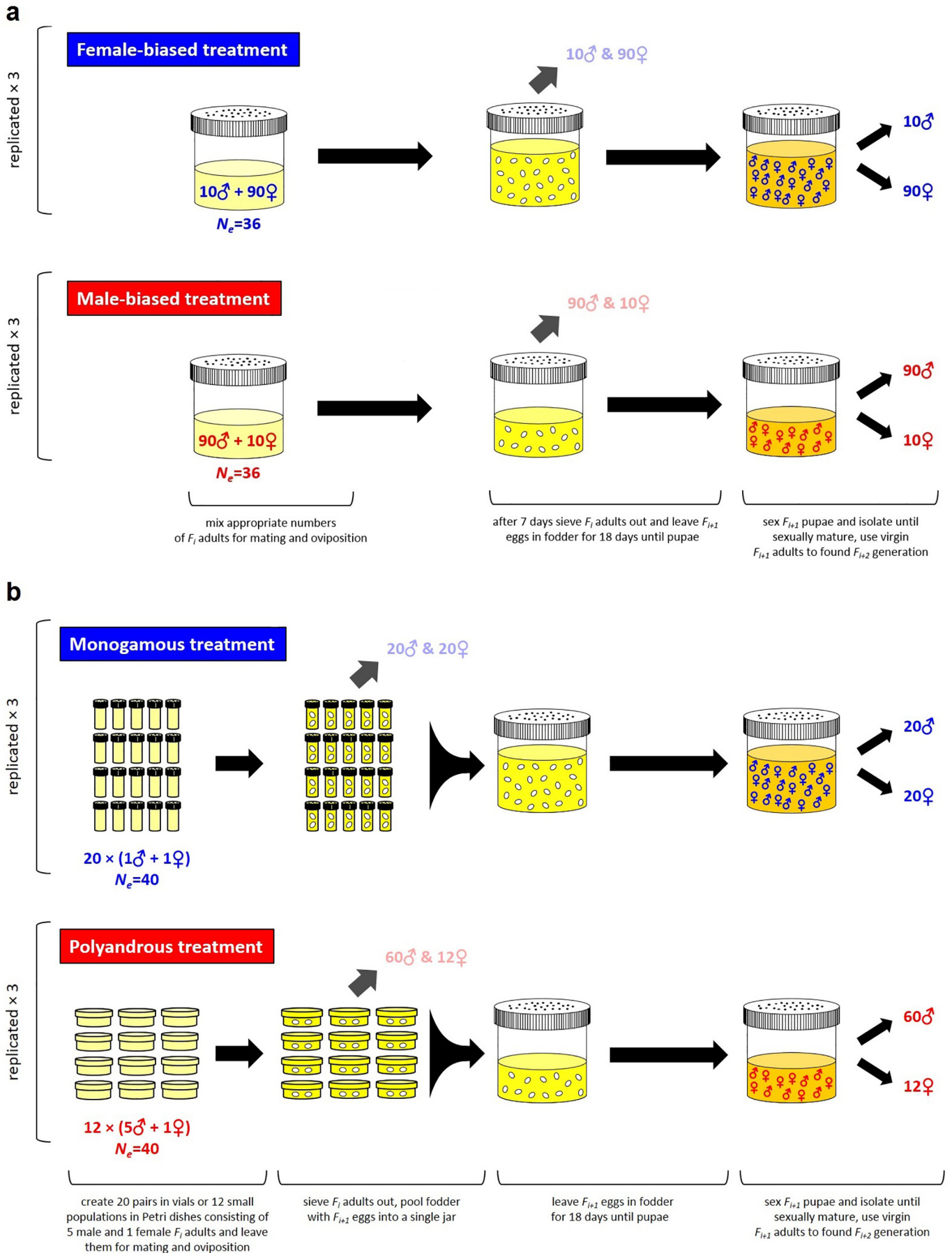
**Code availability.** Scripts of analyses and code used for figure production are available upon request from M.J.G.G.

31. Demont, M. *et al.* Experimental removal of sexual selection reveals adaptations to polyandry in both sexes. *Evol. Biol.* **41**, 62–70 (2014).
32. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* 4th edn (Pearson Education, 1996).
33. Therneau, T. A package for survival analysis in S. R package version 2.37-7. <http://cran.r-project.org/web/packages/survival/index.html> (2014).
34. R Core Team. R: a language and environment for statistical computing (R Foundation for Statistical Computing, 2013).
35. Fournier, D. A. *et al.* AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Meth. Software* **27**, 233–249 (2012).
36. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
37. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-6. <http://CRAN.R-project.org/package=lme4> (2014).
38. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-6. <http://CRAN.R-project.org/package=lmerTest> (2014).



**Extended Data Figure 1 | Experimental rationale for purging and then exposing mutation load.** Having been changed by strong (+SS, red) versus weak (−SS, blue) histories of sexual selection, while under equal influences of natural selection (NS), variation in mutation load residing in the form of recessive alleles is exposed via inbreeding. Inbreeding was enforced through

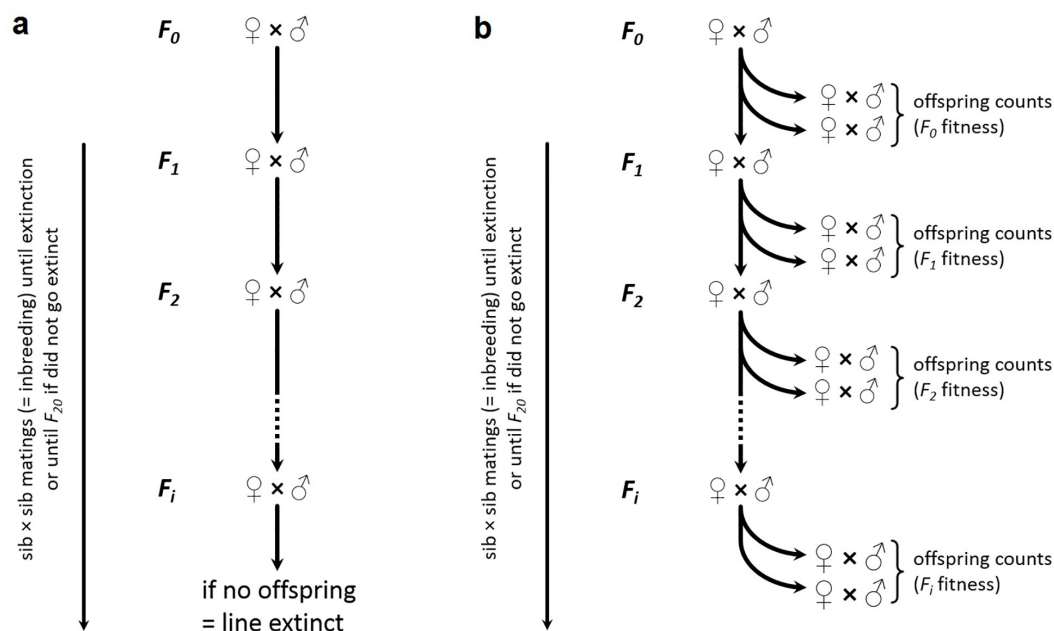
monogamous sib × sib pairings, also eliminating concurrent confounds of interlocus sexual conflict. Populations with reduced mutation load as a result of histories of strong sexual selection are predicted to resist extinction (survival,  $s$ ) and maintain fitness ( $f$ ) under continuous inbreeding ( $i$ ).



**Extended Data Figure 2 | Experimental evolution protocols for regimes A and B.** Contrasting intensities of strong (red) versus weak (blue) sexual selection were imposed upon each generation of adult reproduction, while equalizing effective population size within a regime, and allowing full genetic

mixing within the replicate lines at the egg/larval/pupal stages. From the start, each treatment was replicated to create three independent lines. Regime A (a) applied contrasting sexual selection by varying adult operational sex ratio, while regime B (b) enforced monogamy to compare against polyandry.

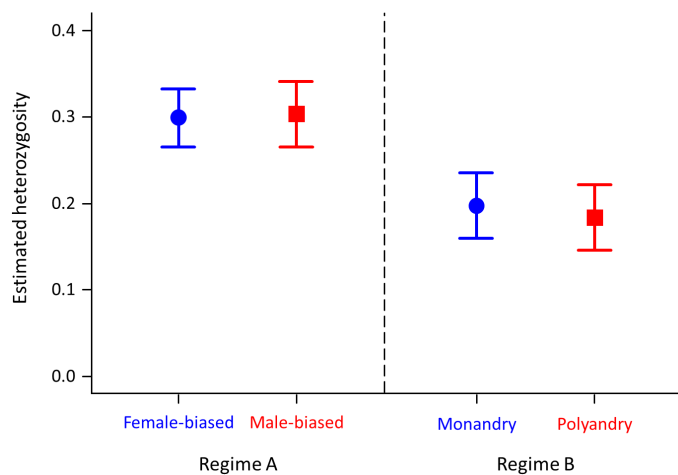




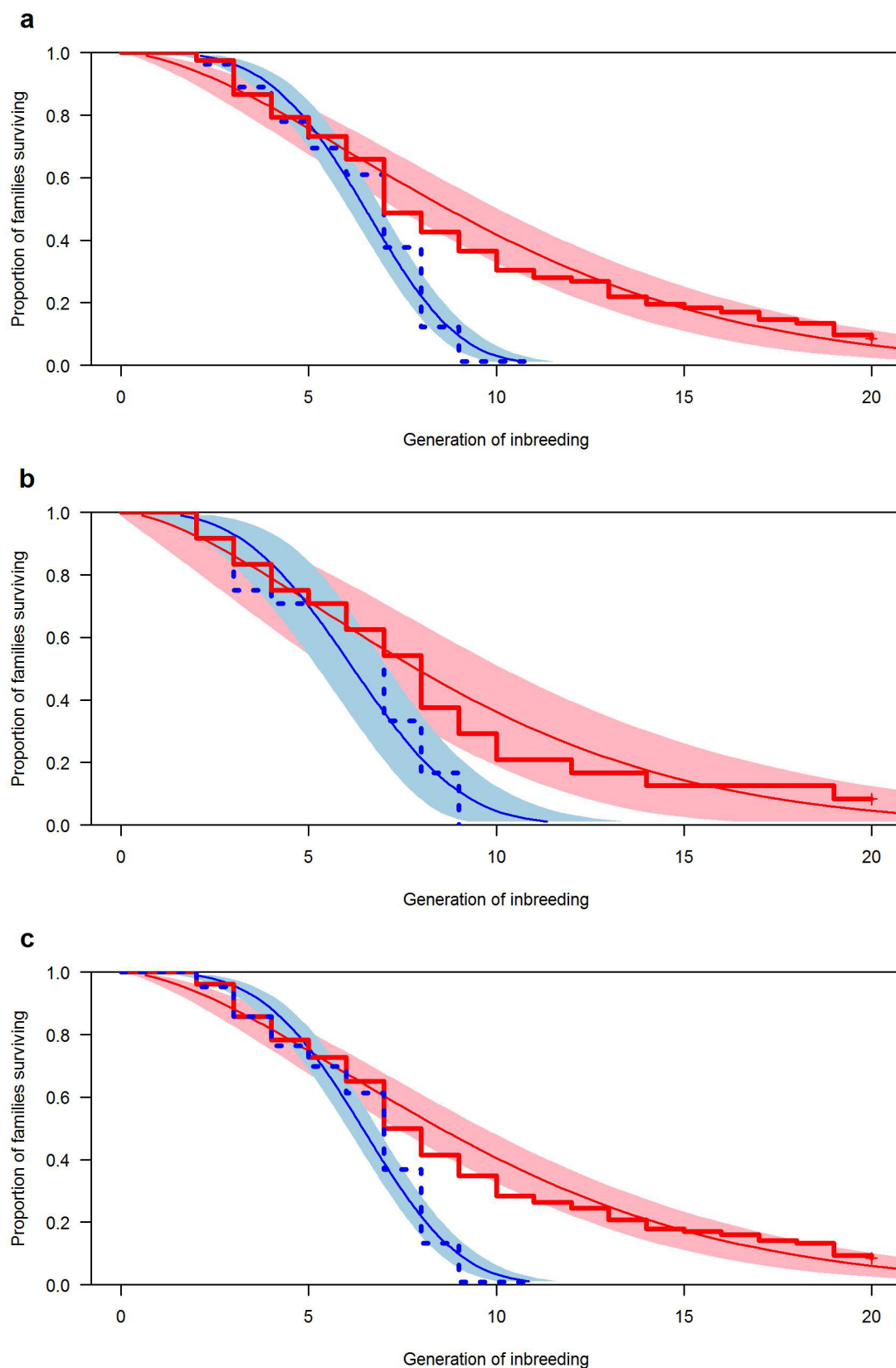
### Extended Data Figure 3 | Extinction and fitness decline protocols.

Inbreeding in family lines was performed via sib  $\times$  sib crosses for up to 20 generations across 3 years. To measure extinction (**a**), a family was considered extinct when it failed to produce offspring, or offspring were of the same sex (which occurred in only 9 out of 216 family lines, indicating no sex-specific pre-adult mortality by treatment). In regime A, extinction data were collected from 28 initial families per line, three lines per sexual selection treatment, comparing both strong versus weak treatments ( $n = 168$  total family lines). In regime B,

extinction data were collected from eight initial families per line, three lines per sexual selection treatment, comparing both strong versus weak treatments ( $n = 48$  total family lines). To measure fitness decline (**b**), two additional sib  $\times$  sib pairs per family per generation were bred to estimate reproductive fitness in every generation by counting number of offspring produced (see Methods). In both regimes A and B, fitness data were collected from eight initial families per line, three lines per sexual selection treatment, and both strong versus weak treatment contrasts in each.

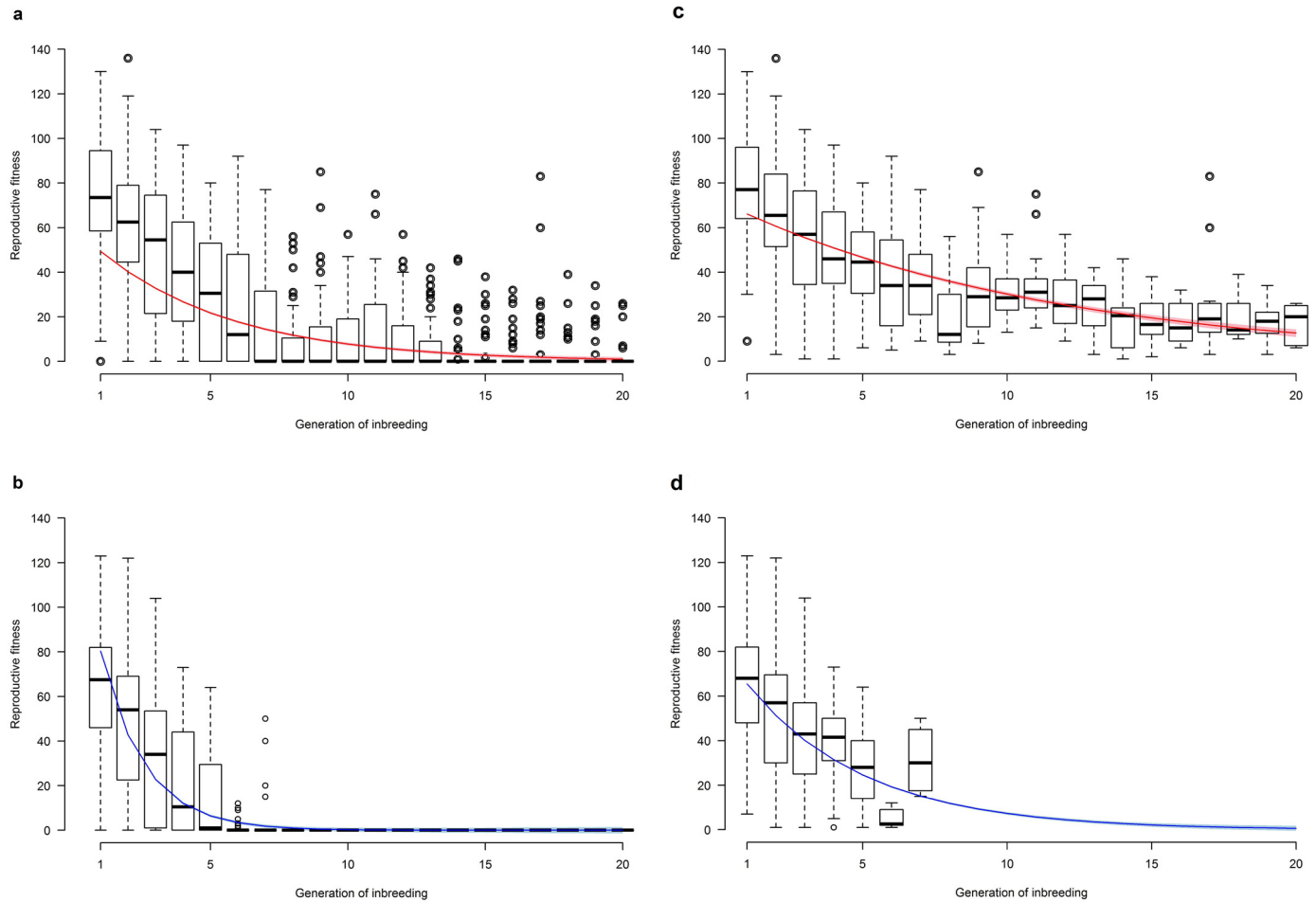


**Extended Data Figure 4 | Estimated heterozygosity ( $\pm$ s.e.m.) does not differ between experimental evolution sexual selection treatments within regime A (left) and regime B (right).** Linear mixed effect modelling showed the estimated heterozygosity of the male-biased selection treatment (M:  $H_{\text{est}} = 0.312$ ,  $t = 9.468$ ) is not significantly different from that of female-biased (F:  $H_{\text{est}} = 0.318$ ,  $t = 9.295$ ,  $P = 0.863$ ), but is significantly different from monogamous and polyandrous treatments (Mo:  $H_{\text{est}} = 0.199$ ,  $t = 6.453$ ,  $P = 0.003$ ; Po:  $H_{\text{est}} = 0.197$ ,  $t = 6.397$ ,  $P = 0.003$ ). The estimated heterozygosities of monogamous and polyandrous treatments are not significantly different ( $P = 0.956$ ) (see Methods).



**Extended Data Figure 5 | Concordance between raw data and model fit in extinction analyses.** Survival curves of raw data (thick and dotted lines) overlaid on model fit (shaded areas with mean curves and 95% confidence intervals). Survival of families derived from strong (red, solid line) or weak (blue, dotted line) sexual selection treatment histories differed: (a) regime A,

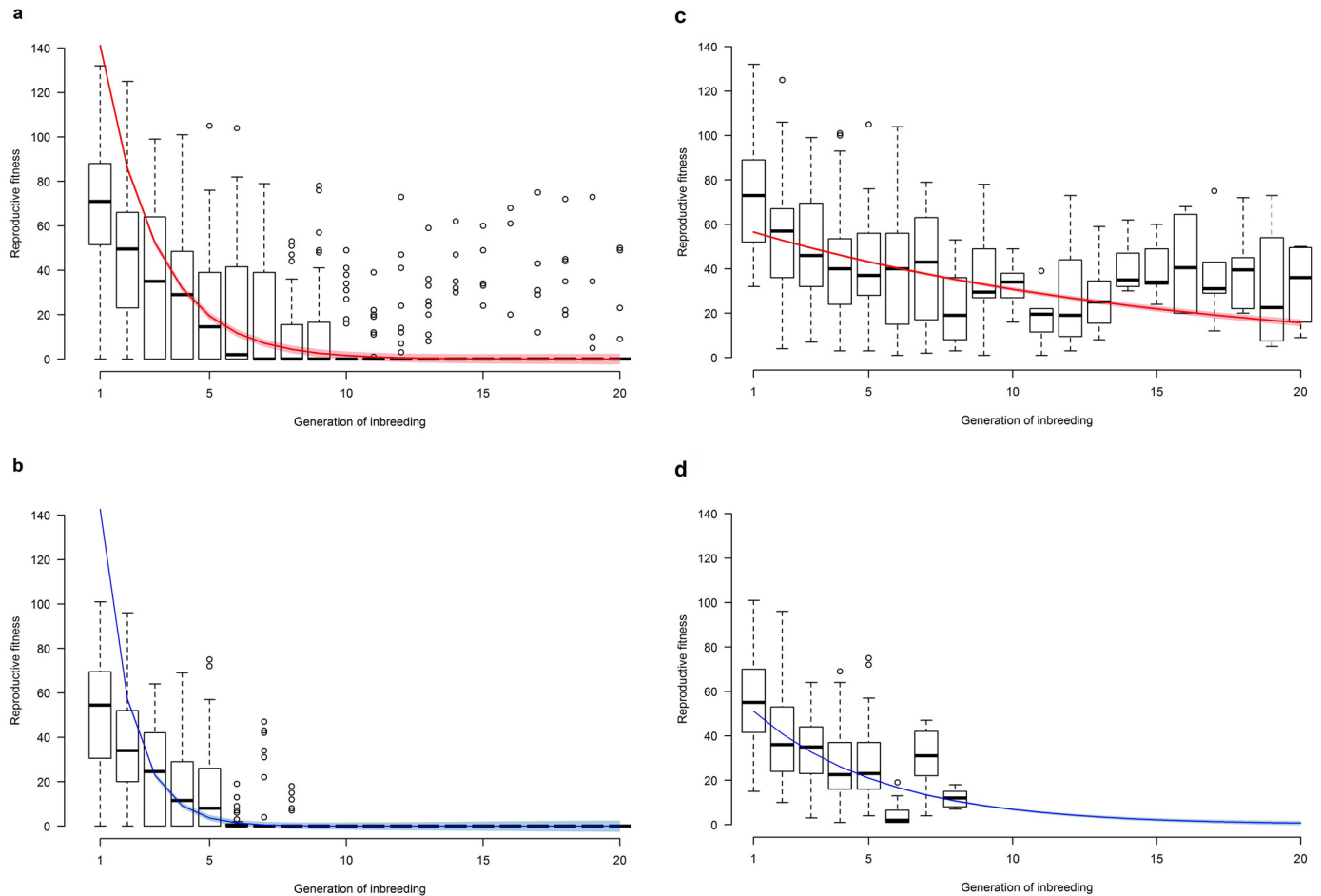
male-biased (red) versus female-biased (blue) sexual selection treatments; (b) regime B, polyandrous (red) versus monogamous (blue); (c) regimes A and B combined into a single analysis. See Fig. 1 and the main text for results of statistical analyses, and Methods and Extended Data Figs 2 and 3 for details of protocols, methods and experimental design.



**Extended Data Figure 6 | Regime A.** Boxplots of the relationships between fitness and inbreeding generation for the male-biased (a and c) versus the female-biased (b and d) treatments. Curves show the predicted relationships between reproductive fitness and inbreeding generation from the GLMMs, and the narrow red and blue shadows show the 95% confidence intervals predicted from the fixed effects. Horizontal bars indicate medians, boxes indicate interquartile ranges, whiskers indicate minimum and maximum values, and

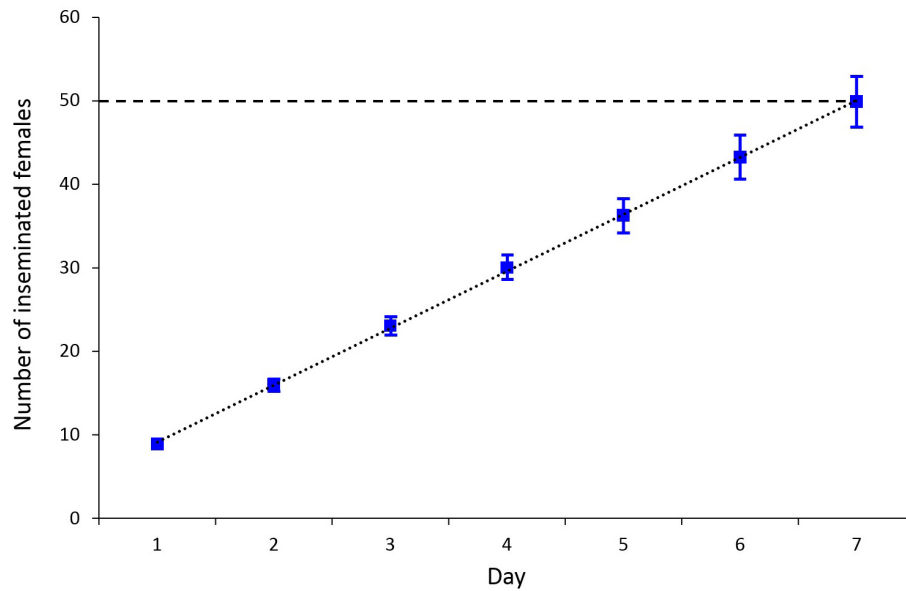
circles indicate outliers (values 1.5 times higher or lower than the first and third quartiles, respectively). Comparison of a versus b identifies the difference in total fitness declines between strong versus weak sexual selection histories in regime A, while c versus d identifies the same difference in decline for fitness but only for the sibling pairs that produced at least some offspring (that is, omitting zero fitness values that may have resulted from a failure to mate). See Fig. 1, main text and Extended Data Table 1 for results of statistical analyses.





**Extended Data Figure 7 | Regime B.** Boxplots of the relationships between fitness and inbreeding generation for the polyandrous (a and c) versus the monogamous (b and d) treatments. Curves show the predicted relationships between reproductive fitness and inbreeding generation from the GLMMs, and the narrow red and blue shadows show the 95% confidence intervals predicted from the fixed effects. Horizontal bars indicate medians, boxes indicate interquartile ranges, whiskers indicate minimum and maximum values, and

circles indicate outliers (values 1.5 times higher or lower than first and third quartiles, respectively). Comparison of a versus b identifies the difference in total fitness declines between strong versus weak sexual selection histories in regime B, while c versus d identifies the same difference in decline for fitness but only for the sibling pairs that produced at least some offspring (that is, omitting zero fitness values that may have resulted from a failure to mate). See Fig. 1, main text and Extended Data Table 1 for results of statistical analyses.



**Extended Data Figure 8 | Across 7 days of mating opportunity, males successfully inseminated 50 females on average ( $\pm$ s.e.m.).** Six virgin females were allocated to individual GA1 control stock males ( $n = 11$ ) every 12 h for 7 days, providing males with 84 potential mates. Over this 1 week period

(replicating that applied within the experimental evolution protocols, Extended Data Fig. 2), males successfully inseminated and generated offspring from an average of 50 females (see Methods).

**Extended Data Table 1 |** Fixed-effect parameter estimates from negative binomial GLMMs of the relationship between fitness and generation of inbreeding for male-biased and female-biased treatments (regime A), and polyandrous and monogamous treatments (regime B), and their statistical interactions. See Methods for details of replication and sample sizes.

| Comparison   |                                    | Estimate | Std. Error | z value | Pr(>  z ) |
|--|------------------------------------|----------|------------|---------|-----------|
| <b>Regime A</b><br>Male-biased<br>vs.<br>Female-biased | Intercept                          | 5.02     | 0.17       | 28.23   | <0.001    |
|  | Inbreeding                         | -0.63    | 0.03       | -21.34  | <0.001    |
|  | Treatment (Female- or Male-biased) | -0.91    | 0.24       | -3.75   | <0.001    |
|  | Inbreeding*Treatment (Male-biased) | 0.42     | 0.03       | 13.82   | <0.001    |
| <b>Regime B</b><br>Polyandrous<br>vs.<br>Monogamous    | Intercept                          | 4.48     | 0.23       | 19.39   | <0.001    |
|  | Inbreeding                         | -0.53    | 0.03       | -21.17  | <0.001    |
|  | Treatment (Monogamy or Polyandry)  | -0.63    | 0.32       | -1.97   | 0.048     |
|  | Inbreeding*Treatment (Polyandry)   | 0.29     | 0.03       | 10.56   | <0.001    |

# PPAR- $\alpha$ and glucocorticoid receptor synergize to promote erythroid progenitor self-renewal

Hsiang-Ying Lee<sup>1\*</sup>, Xiaofei Gao<sup>1\*</sup>, M. Inmaculada Barrasa<sup>1</sup>, Hu Li<sup>2</sup>, Russell R. Elmes<sup>1</sup>, Luanne L. Peters<sup>3</sup> & Harvey F. Lodish<sup>1,4</sup>

Many acute and chronic anaemias, including haemolysis, sepsis and genetic bone marrow failure diseases such as Diamond-Blackfan anaemia, are not treatable with erythropoietin (Epo), because the colony-forming unit erythroid progenitors (CFU-Es) that respond to Epo are either too few in number or are not sensitive enough to Epo to maintain sufficient red blood cell production<sup>1–9</sup>. Treatment of these anaemias requires a drug that acts at an earlier stage of red cell formation and enhances the formation of Epo-sensitive CFU-E progenitors. Recently, we showed that glucocorticoids specifically stimulate self-renewal of an early erythroid progenitor, burst-forming unit erythroid (BFU-E), and increase the production of terminally differentiated erythroid cells<sup>10,11</sup>. Here we show that activation of the peroxisome proliferator-activated receptor  $\alpha$  (PPAR- $\alpha$ ) by the PPAR- $\alpha$  agonists GW7647 and fenofibrate synergizes with the glucocorticoid receptor (GR) to promote BFU-E self-renewal. Over time these agonists greatly increase production of mature red blood cells in cultures of both mouse fetal liver BFU-Es and mobilized human adult CD34<sup>+</sup> peripheral blood progenitors, with a new and effective culture system being used for the human cells that generates normal enucleated reticulocytes. Although *Ppara*<sup>−/−</sup> mice show no haematological difference from wild-type mice in both normal and phenylhydrazine (PHZ)-induced stress erythropoiesis, PPAR- $\alpha$  agonists facilitate recovery of wild-type but not *Ppara*<sup>−/−</sup> mice from PHZ-induced acute haemolytic anaemia. We also show that PPAR- $\alpha$  alleviates anaemia in a mouse model of chronic anaemia. Finally, both in control and corticosteroid-treated BFU-E cells, PPAR- $\alpha$  co-occupies many chromatin sites with GR; when activated by PPAR- $\alpha$  agonists, additional PPAR- $\alpha$  is recruited to GR-adjacent sites and presumably facilitates GR-dependent BFU-E self-renewal. Our discovery of the role of PPAR- $\alpha$  agonists in stimulating self-renewal of early erythroid progenitor cells suggests that the clinically tested PPAR- $\alpha$  agonists we used may improve the efficacy of corticosteroids in treating Epo-resistant anaemias.

The therapeutic effect of glucocorticoids in treating Epo-resistant anaemias such as Diamond-Blackfan anaemia (DBA) is well documented, although steroid therapy has severe side effects that limit its use<sup>12,13</sup>. Given the physiological importance and attractive drug targets of nuclear receptors (NRs)<sup>14–17</sup>, and in the hope of identifying other small-molecule drugs that can either individually or in combination with glucocorticoids expand BFU-E progenitors, we analysed the expression of 49 NR genes during mouse BFU-E self-renewal promoted by dexamethasone (DEX)<sup>10</sup>. Although the expression of most NR genes was unchanged, seven were upregulated and three downregulated by more than twofold (Supplementary Table 1). Among these, PPAR- $\alpha$  changed the most: a fivefold induction (Extended Data Fig. 1a). While most NR agonists and antagonists had no significant effects on erythroid production in cultures of primary mouse BFU-Es, two PPAR- $\alpha$  agonists, GW7647 and fenofibrate, synergized with DEX to increase the

output of erythroid cells (Supplementary Table 2 and Fig. 1a). In the absence of corticosteroids, neither GW7647 nor fenofibrate affected erythropoiesis. Fenofibrate is a US Food and Drug Administration (FDA)-approved drug primarily used to treat hypercholesterolaemia and hypertriglyceridaemia<sup>18</sup>, and GW7647 is a PPAR- $\alpha$  agonist with higher receptor binding affinity and specificity (half-maximum effective concentration (EC<sub>50</sub>) = 6 nM)<sup>19</sup>.

When added together with DEX, either GW7647 or fenofibrate led to a ~150-fold increase in erythroblast production, 3–5-fold greater than DEX alone (Fig. 1a). At the end of the culture, virtually all of the cells were erythroblast cells (Extended Data Fig. 1b). Importantly, the expression of *Ppara* declines markedly from the BFU-E to the CFU-E stage, consistent with the decline of DNase I hypersensitivity of the *Ppara* promoter (Extended Data Fig. 1c, d).

Neither GW7647 nor fenofibrate were able to increase erythroblast production from isolated mouse CFU-E cells, suggesting a specific function of PPAR- $\alpha$  agonists on BFU-Es (Extended Data Fig. 1e). GW7647 also synergized with lower doses of DEX in promoting erythroid cell production (Fig. 1b). Importantly, the effects of GW7647 and fenofibrate require the PPAR- $\alpha$  receptor, since BFU-E cells isolated from *Ppara*<sup>−/−</sup> mice<sup>20</sup> failed to respond to their treatment (Fig. 1a).

Added together with DEX, 10  $\mu$ M of either GW7647 or fenofibrate significantly increased BFU-E colony numbers and maintained high BFU-E numbers over one additional day relative to that achieved by DEX alone (Fig. 1c). Concentrations of GW7647 as low as 100 nM effectively synergized with 100 nM DEX in promoting BFU-E self-renewal (Extended Data Fig. 1f).

Blood from *Ppara*<sup>−/−</sup> mice exhibits no haematological abnormalities, indicating that PPAR- $\alpha$  is not necessary for normal erythropoiesis (Supplementary Table 3). Consistent with this, the PPAR- $\alpha$  antagonist GW6471 does not interfere with DEX-promoted BFU-E self-renewal, indicating that DEX does not promote BFU-E self-renewal through activating the transcriptional activity of PPAR- $\alpha$  (Extended Data Fig. 1g).

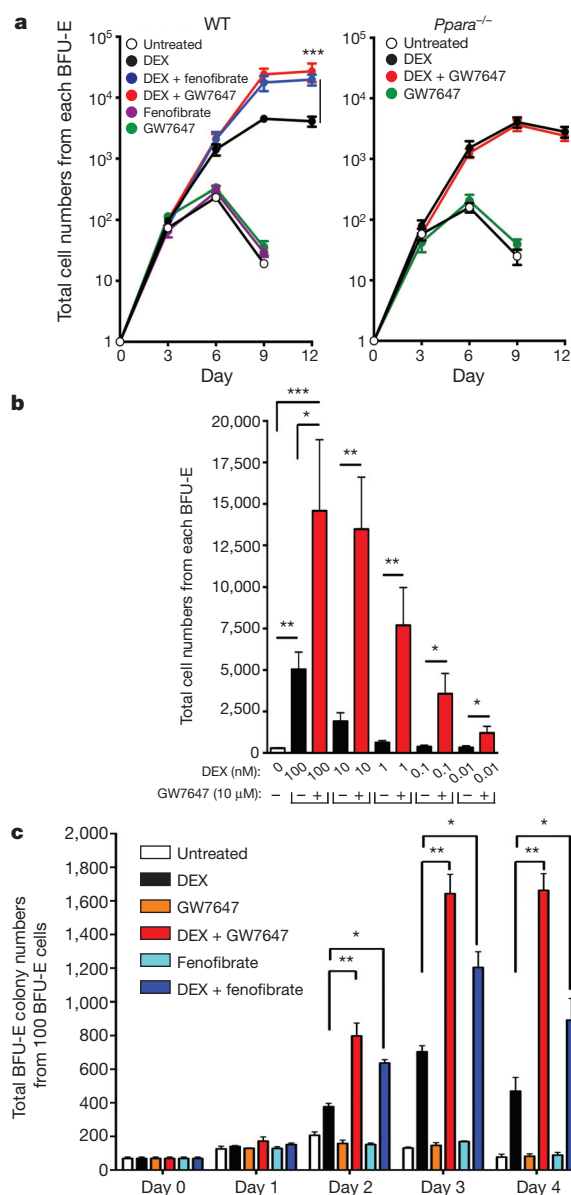
Adding GW7647 increases total cell number by an additional fourfold, yielding a 120,000-fold expansion in our newly developed synchronized human CD34<sup>+</sup> erythroid culture system (Supplementary Discussion and Fig. 2a). In these cultures the numbers of BFU-Es and CFU-Es are highest during days 5–8 (Fig. 2b and Extended Data Fig. 2a). Addition of GW7647 with DEX increased total BFU-E numbers, which in turn subsequently increased total CFU-E numbers in the culture. Similar to mouse BFU-E cultures, GW7647 synergized with very low concentrations of DEX in stimulating erythroid expansion of human CD34<sup>+</sup> cells (Extended Data Fig. 2b). Addition of DEX and GW7647 affects expression of many key erythroid genes (Supplementary Discussion).

As expected, knocking down PPAR- $\alpha$  abrogated the ability of GW7647 to stimulate production of BFU-E cells or the total number of erythroid cells at the end of the culture (Fig. 2c and Extended

<sup>1</sup>Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. <sup>2</sup>Center for Individualized Medicine, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, Minnesota 55905, USA. <sup>3</sup>The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609, USA. <sup>4</sup>Departments of Biology and Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

\*These authors contributed equally to this work.

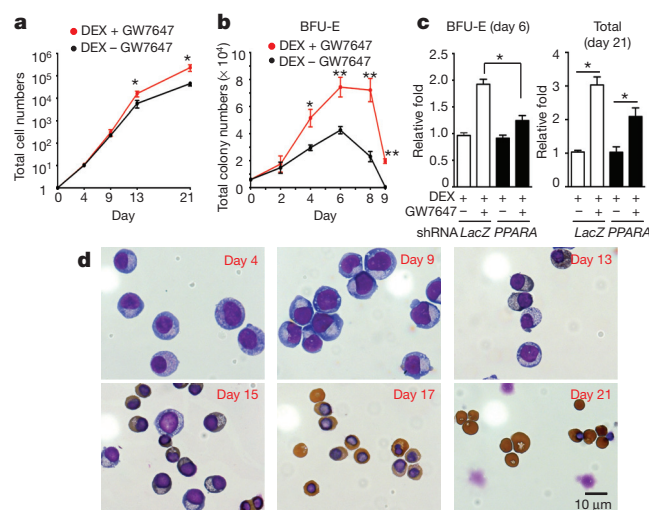




**Figure 1 | PPAR- $\alpha$  signalling synergizes with GR to promote self-renewal of mouse BFU-E erythroid progenitors.** **a**, Wild-type (WT; left) and *Ppara*<sup>-/-</sup> (right) BFU-E cells from embryonic day (E) 14.5 mouse fetal livers were isolated and cultured in serum-free erythroid liquid expansion (SFELE) medium with the indicated treatment. Cell numbers were counted every 3 days. **b**, Mouse BFU-E cells were cultured in DEX ± 10 μM GW7647 as indicated. Total erythroid cell numbers were counted at day 9. **c**, Colony-forming assays were conducted to determine BFU-E colony numbers from 100 mouse BFU-E cells cultured under the indicated conditions. Colony-forming assays were performed at 24-h intervals. Data are mean ± s.d. from three biological replicates; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, Student's *t*-test.

Data Fig. 2c). Consistent with the absence of a role for PPAR- $\alpha$  in normal haematopoiesis, knocking down PPAR- $\alpha$  in human CD34<sup>+</sup> cells did not affect the ability of DEX to stimulate BFU-E production or production of erythroid cells. Thus, in the human as well as in the mouse, only ligand-activated, not unactivated, PPAR- $\alpha$  synergizes with GR to promote BFU-E self-renewal during erythroid differentiation. We note that our cultures are highly synchronous on the basis of cell surface marker expression (Extended Data Fig. 2e) as well as morphology (Fig. 2d). Neither DEX nor GW7647 had any effects on terminal differentiation or enucleation (Extended Data Fig. 2h).

GW7647 significantly increased the number of both BFU-E and CFU-Es in our CD34<sup>+</sup> cell culture system after RPS19 knockdown, which



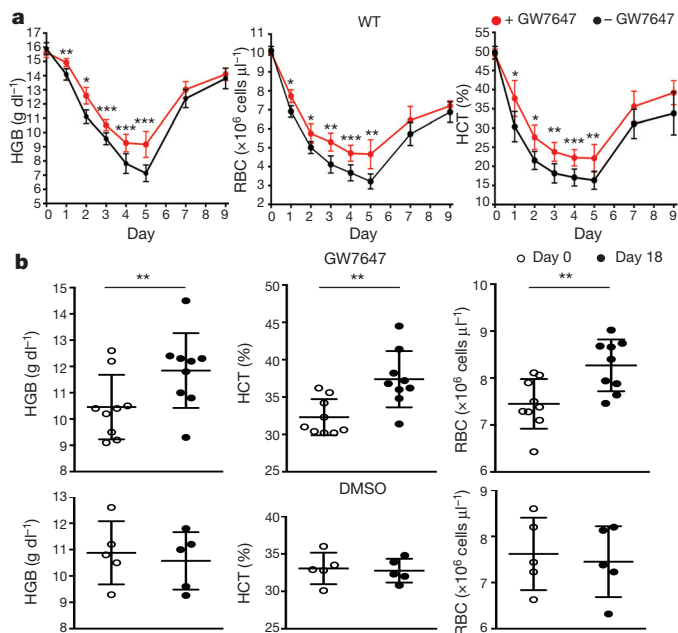
**Figure 2 | Activation of PPAR- $\alpha$  signalling increased erythroid cell expansion in the *ex vivo* human CD34<sup>+</sup> erythroid culture system.** **a**, Human CD34<sup>+</sup> cells were cultured as described in Methods. Total cell numbers were quantified. **b**, BFU-E colony numbers were quantified by plating 1,000 cells at various time points during days 0–9 of the human CD34<sup>+</sup> erythroid culture on methylcellulose. **c**, BFU-E numbers (left) or total cell number (right) from control or PPAR- $\alpha$ -knockdown human CD34<sup>+</sup> cells cultured under the indicated condition were counted. shRNA, short hairpin RNA. **d**, Benzidine-Giemsa staining demonstrating cell morphology of the *ex vivo* human CD34<sup>+</sup> erythroid differentiation system. \**P* < 0.05, \*\**P* < 0.01, Student's *t*-test. Data are mean ± s.d. from three biological replicates.

recapitulates RPS19 haploinsufficiency in DBA patients<sup>21,22</sup> (Extended Data Fig. 3a). Addition of GW7647 substantially increased the fraction of CD71<sup>+</sup> cells at day 9 of culture after RPS19 knockdown as well as total cell numbers and the fraction of CD235a<sup>+</sup> cells at day 21 (Extended Data Figs 3b, c). Taken together, our data suggest that GW7647 facilitates BFU-E self-renewal and production of immature erythroid progenitors in both wild-type and RPS19-knockdown human cells.

We next tested the function of GW7647 in a PHZ-induced haemolytic anaemia mouse model in which the endogenous corticosteroid level becomes markedly increased<sup>11</sup>. Wild-type mice were treated with either dimethylsulfoxide (DMSO) or GW7647 for 3 days followed by treatment with PHZ; they were then injected with DMSO or GW7647 for another 7 days (Supplementary Discussion and Extended Data Fig. 4a). Treatment with GW7647 resulted in significantly higher levels of haemoglobin, red blood cell numbers, and haematocrit after PHZ injection compared to in control mice (Fig. 3a). In contrast, white blood cell counts were similar in the two groups (data not shown). Importantly, the function of GW7647 during stress erythropoiesis was dependent on PPAR- $\alpha$ , as *Ppara*<sup>-/-</sup> mice failed to respond to GW7647 treatment (Extended Data Fig. 4b).

We also tested the ability of PPAR- $\alpha$  agonists to stimulate red cell production in a mouse model of chronic anaemia, 'neonatal anaemia' (*Nan*) mice (Supplementary Discussion). While BFU-E numbers in spleens of *Nan*/+ mutant mice are similar to or slightly higher than that in wild-type mice, the average number of BFU-Es in the spleens of GW7647-injected *Nan*/+ mutant mice are higher than that in untreated mice (Extended Data Fig. 5b). Importantly, GW7647 injection increased haemoglobin level, haematocrits and red blood cell numbers in these anaemic mice (Fig. 3b), suggesting that the PPAR- $\alpha$  agonist alleviates anaemia in *Nan*/+ mutant mice by increasing BFU-E numbers in the spleen. GW7647 does not have any significant effects on either platelet or white blood cell numbers in peripheral blood from *Nan*/+ mice (Extended Data Fig. 5c).

To understand the molecular mechanism by which PPAR- $\alpha$  synergizes with GR to promote BFU-E self-renewal, we conducted chromatin immunoprecipitation followed by sequencing (ChIP-seq)

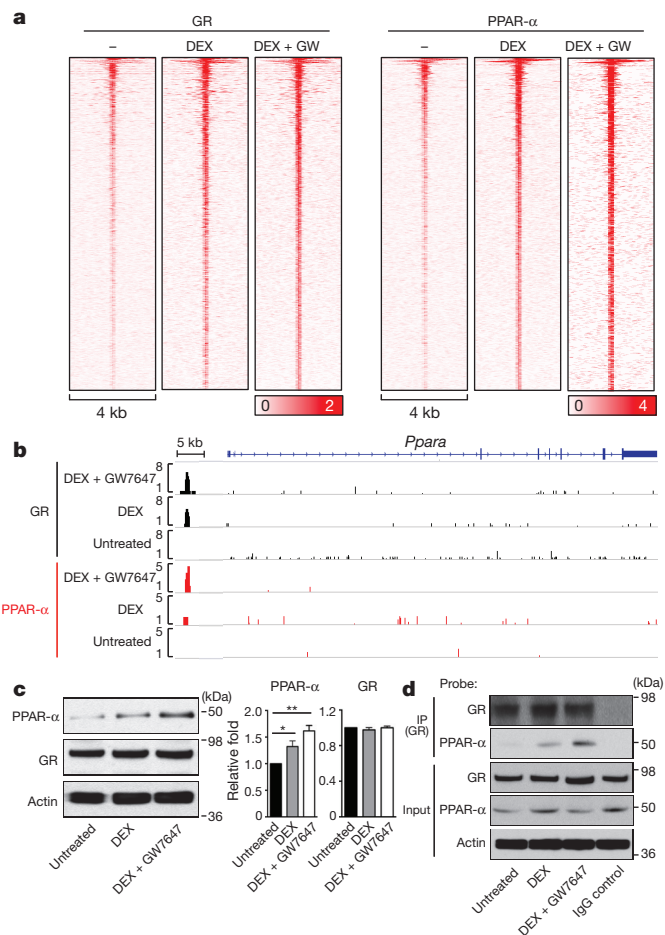


**Figure 3** | PPAR- $\alpha$  agonist GW7647 is effective *in vivo* to alleviate anaemic symptoms. **a**, Wild-type mice were pretreated with DMSO or 100  $\mu$ g kg<sup>-1</sup> GW7647 for 3 days (days -3 to -1) before PHZ injection on day 0. Mice were subject to daily DMSO or GW7647 injections during days 0–6. Haemoglobin (HGB), red blood cells (RBC) and haematocrit (HCT) were measured in wild-type mice on the days indicated.  $n = 6$ . **b**, *Nan*<sup>+/+</sup> mutant mice were injected with DMSO or GW7647 (100  $\mu$ g kg<sup>-1</sup>) for 18 days. Each dot represents one mouse. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , Student's *t*-test.

analysis to interrogate the genome-wide chromatin occupancy of GR and PPAR- $\alpha$  in mouse BFU-E cells. Compared to untreated cells, GR and PPAR- $\alpha$  occupancy were both induced upon DEX treatment alone; addition of GW7647 further enhanced PPAR- $\alpha$  but not GR occupancy (Fig. 4a). These results indicate that GW7647 enhances the recruitment of PPAR- $\alpha$  to GR-binding sites in BFU-Es. In BFU-Es co-treated with GW7647 and DEX, our ChIP-seq detected 1,058 GR and 1,623 PPAR- $\alpha$  peaks. GR and PPAR- $\alpha$  both predominantly occupied distal intergenic (>63%) and intronic chromatin sites (>25%) (Extended Data Fig. 6a). Notably, in 719 peaks, GR and PPAR- $\alpha$  localize in close proximity (Extended Data Fig. 6b); 67.9% of total GR peaks and 44.3% of PPAR- $\alpha$  peaks co-localize. The DNA sequences underlying these overlapping peaks are enriched for DNA-binding motifs for several transcription factors including PU.1, YY1, Smad3, Tal1, Klf4, Hif-2 $\alpha$  and Myb (Extended Data Fig. 6c); most of these transcription factors are known to have crucial roles in stem-cell self-renewal<sup>23</sup>. In addition, co-treatment with DEX and GW7647 also upregulates many genes critical for stem-cell self-renewal (Supplementary Discussion and Extended Data Figs 7, 8).

DEX treatment of BFU-E cells leads to a slight increase in PPAR- $\alpha$  protein levels, and this is further increased to 1.6 times that of control cells by treatment with DEX and GW7647 (Fig. 4c). In contrast, the GR protein level is not altered in BFU-Es under these conditions. Consistent with these observations, addition of DEX to BFU-E cells induced the binding of both GR and PPAR- $\alpha$  to a chromatin site ~5 kb upstream of the transcription start site of *Ppara*, presumably part of a *Ppara* gene enhancer, and the binding of PPAR- $\alpha$  to this chromatin site was further enhanced by GW7647 addition (Fig. 4b). Our data suggest that there is a positive autoregulatory feedback loop of PPAR- $\alpha$  expression during BFU-E self-renewal promoted by DEX and GW7647.

Co-immunoprecipitation experiments demonstrated an interaction of GR and PPAR- $\alpha$  only in BFU-E cells treated with DEX with or without GW7647 (Fig. 4d). While addition of DEX to cultures of BFU-E cells stimulated PPAR- $\alpha$  binding to the GR, interactions between



**Figure 4** | PPAR- $\alpha$  and glucocorticoid signalling pathways regulate a common target gene ensemble. **a**, Density maps of GR and PPAR- $\alpha$  ChIP-seq signals in mouse BFU-E cells. GR-binding peaks in DEX and GW7647 co-treated samples were used as the reference to search for corresponding ChIP-seq signals of PPAR- $\alpha$ . **b**, GR and PPAR- $\alpha$  occupancy across the *Ppara* locus in BFU-E cells treated as indicated. **c**, Left, protein level of GR, PPAR- $\alpha$  and  $\beta$ -actin in mouse BFU-E cells treated as indicated. Right, densitometry quantification of GR and PPAR- $\alpha$  protein expression levels normalized to  $\beta$ -actin. Experiments were repeated three times. \* $P < 0.05$ , \*\* $P < 0.01$ , Student's *t*-test. **d**, Co-immunoprecipitation (IP) demonstrating interaction between GR and PPAR- $\alpha$  in mouse BFU-E cells.

PPAR- $\alpha$  and GR were more pronounced after treatment with both DEX and GW7647. These results extend our ChIP-seq data, indicating that there is a physical interaction between GR and PPAR- $\alpha$  and probably other proteins that underlies DEX- and GW7647-enhanced BFU-E self-renewal (Supplementary Discussion and Extended Data Fig. 9).

Given that little is known concerning endogenous PPAR- $\alpha$  ligands, the precise role of PPAR- $\alpha$  in haematopoiesis, and erythropoiesis in particular, remains elusive. Nonetheless, our finding that agonists of PPAR- $\alpha$  lead to enhanced binding of PPAR- $\alpha$  to many chromatin sites and to an increase in corticosteroid-induced BFU-E self-renewal and erythroid expansion, point to a previously unappreciated role for this nuclear receptor in haematopoiesis (Extended Data Fig. 10). The function of PPAR- $\alpha$  has been mainly studied in nutrient metabolism and energy homeostasis, and fenofibrate is already an FDA-approved drug for dyslipidaemia treatment<sup>18</sup>. Given that there is a very limited number of drugs that can be used to treat Epo-resistant anaemias, the development of new drugs or repurposing current drugs to treat these diseases is challenging. Our surprising discovery suggests a novel function of PPAR- $\alpha$  in self-renewal of early committed erythroid progenitors, which could potentially lead to new therapeutics to treat Epo-resistant anaemias such as DBA.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 8 April 2014; accepted 12 February 2015.**

**Published online 11 May 2015.**

1. Kaushansky K. *et al.* *Williams Hematology* 8th edn (McGraw Hill, 2010).
2. Livingston, D. H. *et al.* Bone marrow failure following severe injury in humans. *Ann. Surg.* **238**, 748–753 (2003).
3. Jones, K. B., Anderson, D. W. & Longmore, G. D. Effects of recombinant hematopoietins on blood-loss anemia in mice. *Iowa Orthop. J.* **25**, 129–134 (2005).
4. Robinson, Y. *et al.* Impaired erythropoiesis after haemorrhagic shock in mice is associated with erythroid progenitor apoptosis *in vivo*. *Acta Anaesthesiol. Scand.* **52**, 605–613 (2008).
5. Zimmerman, J. L. Use of blood products in sepsis: an evidence-based review. *Crit. Care Med.* **32**, S542–S547 (2004).
6. Fiorillo, A. *et al.* Unresponsiveness to erythropoietin therapy in a case of Blackfan Diamond anemia. *Am. J. Hematol.* **37**, 65 (1991).
7. Nathan, D. G., Clarke, B. J., Hillman, D. G., Alter, B. P. & Housman, D. E. Erythroid precursors in congenital hypoplastic (Diamond-Blackfan) anemia. *J. Clin. Invest.* **61**, 489–498 (1978).
8. Vlachos, A. & Muir, E. How I treat Diamond-Blackfan anemia. *Blood* **116**, 3715–3723 (2010).
9. Niemeyer, C. M. *et al.* Treatment trial with recombinant human erythropoietin in children with congenital hypoplastic anemia. *Contrib. Nephrol.* **88**, 276–280 (1991).
10. Flygare, J., Rayon Estrada, V., Shin, C., Gupta, S. & Lodish, H. F. HIF1 $\alpha$  synergizes with glucocorticoids to promote BFU-E progenitor self-renewal. *Blood* **117**, 3435–3444 (2011).
11. Zhang, L. *et al.* ZFP36L2 is required for self-renewal of early burst-forming unit erythroid progenitors. *Nature* **499**, 92–96 (2013).
12. Willig, T. N. *et al.* Identification of new prognosis factors from the clinical and epidemiologic analysis of a registry of 229 Diamond-Blackfan anemia patients. *Pediatr. Res.* **46**, 553–561 (1999).
13. Lipton, J. M., Atsidaftos, E., Zyskind, I. & Vlachos, A. Improving clinical care and elucidating the pathophysiology of Diamond Blackfan anemia: an update from the Diamond Blackfan Anemia Registry. *Pediatr. Blood Cancer* **46**, 558–564 (2006).
14. Bauer, A. *et al.* The glucocorticoid receptor is required for stress erythropoiesis. *Genes Dev.* **13**, 2996–3002 (1999).
15. von Lindern, M. *et al.* The glucocorticoid receptor cooperates with the erythropoietin receptor and c-Kit to enhance and sustain proliferation of erythroid progenitors *in vitro*. *Blood* **94**, 550–559 (1999).
16. O'Malley, B. The steroid receptor superfamily: more excitement predicted for the future. *Mol. Endocrinol.* **4**, 363–369 (1990).
17. Moore, J. T., Collins, J. L. & Pearce, K. H. The nuclear receptor superfamily and drug discovery. *ChemMedChem* **1**, 504–523 (2006).
18. Yang, L. P. & Keating, G. M. Fenofibric acid: in combination therapy in the treatment of mixed dyslipidemia. *Am. J. Cardiovasc. Drugs* **9**, 401–409 (2009).
19. Muoio, D. M. *et al.* Peroxisome proliferator-activated receptor- $\alpha$  regulates fatty acid utilization in primary human skeletal muscle cells. *Diabetes* **51**, 901–909 (2002).
20. Lee, S. S. *et al.* Targeted disruption of the alpha isoform of the peroxisome proliferator-activated receptor gene in mice results in abolishment of the pleiotropic effects of peroxisome proliferators. *Mol. Cell. Biol.* **15**, 3012–3022 (1995).
21. Flygare, J. *et al.* Deficiency of ribosomal protein S19 in CD34<sup>+</sup> cells generated by siRNA blocks erythroid development and mimics defects seen in Diamond-Blackfan anemia. *Blood* **105**, 4627–4634 (2005).
22. Ebert, B. L. *et al.* An RNA interference model of RPS19 deficiency in Diamond-Blackfan anemia recapitulates defective hematopoiesis and rescue by dexamethasone: identification of dexamethasone-responsive genes by microarray. *Blood* **105**, 4620–4626 (2005).
23. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
24. Wong, P. *et al.* Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. *Blood* **118**, e128–e138 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank the Whitehead Institute Flow Cytometry Facility, Genome Technology Core and Bioinformatics & Research Computing Facility, as well as the Massachusetts Institute of Technology Koch Institute Flow Cytometry Core. We thank V. Sankaran for haemoglobin HPLC and J. Flygare for the plasmid encoding the RPS19 shRNA, and are grateful to animal technicians F. Reinhardt and T. E. Chavarria for their assistance. We thank T. DiCesare for assistance with graphics. This study was supported by grants to H.F.L. (Defense Advanced Research Projects Agency HR0011-14-2-0005; Department of Defense/US Army Medical Research and Materiel Command W81WH-12-1-0449, National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute 2 P01 HL032262-25; as well as research support from the Diamond-Blackfan Anemia Foundation and Diamond Blackfan Anemia Canada. L.L.P. was supported by NIH grant DK100692. X.G. was supported by a postdoctoral fellowship from the Leukemia and Lymphoma Society.

**Author Contributions** H.-Y.L., X.G., L.L.P. and H.F.L. designed the experiments. H.-Y.L., X.G. and R.R.E. performed the experiments. M.I.B. and H.L. conducted bioinformatic analyses of ChIP-seq and RNA-seq. H.-Y.L., X.G. and H.F.L. wrote the manuscript with input from M.I.B. All authors discussed the results and commented on the manuscript.

**Author Information** RNA-seq and ChIP-seq data have been deposited in the Gene Expression Omnibus under accession numbers GSE63836 and GSE63837, respectively. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.F.L. ([lodish@wi.mit.edu](mailto:lodish@wi.mit.edu)).



## METHODS

**Reagents.** Chemicals were obtained from Sigma. Human peripheral blood granulocyte-colony stimulating factor (G-CSF)-mobilized haematopoietic stem/progenitor cells enriched for CD34<sup>+</sup> were purchased from the Fred Hutchinson Cancer Research Center (FHCRC). StemSpan SFEM, CC100 cytokine cocktail, fetal bovine serum and BSA were purchased from STEMCELL Technologies. Holo human transferrin was from Sigma. Recombinant human and murine stem cell factor (rhSCF, rmSCF) and interleukin-3 (rhIL-3, rmIL-3), recombinant murine interleukin-6 (rmIL-6) and recombinant murine insulin like growth factor-1 (rmIGF-1) were from Peprotech. Recombinant human erythropoietin (rhEpo) was from Amgen. Antibodies used were as follows. Santa Cruz: GR: H300 (sc-8992), M-20 (sc-1004); PPAR- $\alpha$ : H-98 (sc-9000).  $\beta$ -actin: N-21(sc-130656); RPS19 antibody is from Abcam (ab57643). FACS antibodies are from eBioscience: anti-human c-kit PE (#12-1178), anti-human CD71 FITC (#11-0719), anti-human CD235a APC (#17-9987), anti-mouse CD71 PE (#12-0711), and anti-mouse Ter119 APC (#17-5921). Real-time PCR primers (mouse): Hbb-F, 5'-TTTAACGATGGCTGAATCACTT-3'; Hbb-R, 5'-CAGCACAAATCAGATCATATTGC-3'; Slc4a1-F, 5'-GGACAGATAGCATATAGAGACCTAACCA-3'; Slc4a1-R, 5'-CGTAGTCTGTGGCTGTTTGCTC-3'; Egl2-F, 5'-CTGGGCAACTACGTCATCAAT-3'; Egl2-R, 5'-CCGCCATGCACCTTAACATC-3'; Hmgs2-F, 5'-GAAGAGAGCGATGCAGGAAAC-3'; Hmgs2-R, 5'-GTCCACATATTGGGCTGGAAA-3'; Kit-F, 5'-GTTCTGCTCCTACTGCTTCGC-3'; Kit-R, 5'-TAA CAGCCTAATCTCGTCGCC-3'.

shRNA sequences: human PPARA shRNA-1, 5'-GGAGTTTATGAGGCCA TATTC-3'; human PPARA shRNA-2, 5'-GCTTTACGGAATACCACTATT-3'; mouse Pu.1 shRNA, 5'-CCATGTCCACAACAACGAGTT-3'. ChIP-PCR primers: Kit ChIP-PCR-F, 5'-GTCACAGCCACCAGAGAGAG-3'; Kit ChIP-PCR-R, 5'-TGGCAATGTTAAGAAGTGGTGG-3'; Ppara ChIP-PCR-F, 5'-CC AGGCTACACAGAGAAAC-3'; Ppara ChIP-PCR-R, 5'-TTGGCACAATCT CAGGCTA-3'.

RPS19 shRNA is described in a previously published study<sup>21</sup>.

**Ex vivo human CD34<sup>+</sup> erythroid culture.** Our human CD34<sup>+</sup> cell erythroid differentiation method is composed of 4 phases over 21 days: Expansion (day 0–4), differentiation (Dif) I (day 5–9), II (day 10–13), and III (day 14–21). Cells are thawed according to the FHCRC protocol, and cultured in expansion medium (StemSpan SFEM, CC100 cytokine cocktail and 2% penicillin–streptomycin) at  $10^5$  cells ml<sup>-1</sup> from day 0 to 4. After expansion, cells are subsequently cultured in Iscove's modified Dulbecco's medium (IMDM)-based erythroid differentiation medium supplemented with different cytokines in Dif I, II and III. The medium base for all three differentiation phases comprises: IMDM, 15% FBS, 2 mM glutamine, 1% BSA, 500  $\mu$ g ml<sup>-1</sup> holo human transferrin, 10  $\mu$ g ml<sup>-1</sup> recombinant human insulin and 2% penicillin–streptomycin. On days 5–9, cells are cultured in Dif I medium, which contains erythroid medium base, 1  $\mu$ M DEX, 1  $\mu$ M  $\beta$ -oestradiol, 5 ng ml<sup>-1</sup> rhIL-3, 100 ng ml<sup>-1</sup> rhSCF and 6 U ml<sup>-1</sup> rhEpo. On days 10–13, cells are grown in Dif II medium containing erythroid medium base, 50 ng ml<sup>-1</sup> rhSCF and 6 U ml<sup>-1</sup> rhEpo. On days 14–21, cells are cultured in fibronectin-coated plates in Dif III medium, which is erythroid medium base supplemented with 2 U ml<sup>-1</sup> rhEpo. GW7647 was added in both Dif I (100 nM) and Dif II (10 nM) when indicated. Cell numbers re-seeded in the beginning of Dif I, II and III were  $10^5$ ,  $2 \times 10^5$ , and  $3 \times 10^5$  per ml, respectively.

**Mouse fetal liver BFU-E and CFU-E isolation and culture.** Purification of BFU-E and CFU-E cells from murine fetal livers was performed as described previously<sup>10</sup>. Briefly, fetal liver erythroid cells were isolated between E14.5 and E15.5. A pure erythroid progenitor population containing more than 90% of BFU-E and CFU-E cells was obtained through a negative magnetic bead selection for lineage markers including Ter119, B220, Mac-1, CD3, Gr-1, Sca-1, CD16/CD32, CD41, CD34 and positive selection for c-Kit. Afterwards c-Kit<sup>+</sup> BFU-E (CD71 10%<sup>low</sup>) and CFU-E (CD71 20%<sup>high</sup>) cells were separated by flow cytometry. Purified BFU-E or CFU-E cells were seeded in serum-free erythroid liquid expansion medium (SFELE) (StemSpan SFEM with 100 ng ml<sup>-1</sup> rmSCF, 40 ng ml<sup>-1</sup> rmIGF-1, and 2 U ml<sup>-1</sup> rhEpo) alone, or SFELE containing 100 nM DEX  $\pm$  10  $\mu$ M GW7647. Erythroblasts produced from purified BFU-E and CFU-E were analysed over time by FACS and benzidine–Giemsa staining as described<sup>10</sup>.

For Pu.1 knockdown, primary BFU-Es purified from mouse E14.5 fetal liver were incubated with virus encoding a shRNA for LacZ or mouse Pu.1 at 37 °C overnight. After incubation, cells were cultured in SFELE medium with or without DEX and/or GW7647.

**Colony-forming assay.** Murine BFU-E colony-forming assays were performed in MethoCult M3234 (StemCell Technologies) containing 10 U ml<sup>-1</sup> rhEpo, 20 ng ml<sup>-1</sup> rmIL-3, 20 ng ml<sup>-1</sup> rmIL-6 and 50 ng ml<sup>-1</sup> rmSCF, with or without 100 nM DEX  $\pm$  10  $\mu$ M GW7647. For CFU-E colony-forming assays, cells were cultured

in MethoCult M3234 containing 10 U ml<sup>-1</sup> rhEpo, with or without 100 nM DEX  $\pm$  10  $\mu$ M GW7647. The number of CFU-E or BFU-E colonies was scored after 3 days or 7 days in culture, respectively.

For human colony-forming assays, cells were plated in MethoCult H4034 Optimum (StemCell Technologies) without additional cytokines. Cells were cultured for 12–14 days before BFU-E and CFU-E colonies were scored.

**In vivo mice experiments to test PPAR- $\alpha$  agonists.** 129/SvImJ (wild-type), 129S4/SvJae (wild-type), or 129S4/SvJae-Ppara<sup>tm1Gonz/J</sup> female mice (The Jackson Laboratory), 6–8 weeks of age, randomized by weight, were pretreated with GW7647 (100  $\mu$ g kg<sup>-1</sup>) for 3 days (days –3 to –1). On day 0, the mice were injected with phenylhydrazine (PHZ) (60 mg kg<sup>-1</sup>). GW7647 injection was continued during days 0–6. Whole blood samples were collected at each day of day 1–5, and day 7 and 9 for complete blood count (CBC) analyses.

Four- to six-week-old *Nan*<sup>+</sup> mutant mice (officially designated as *Klf1*<sup>Nan</sup>, <http://www.informatics.jax.org/allele/MGI:1861107>), obtained from the Jackson Laboratory, were injected with GW7647 (100  $\mu$ g kg<sup>-1</sup>) for 18 days. Mice were randomized by weight. Whole blood samples were collected every 3 days for CBC analyses.

All mouse procedures were approved by the Animal Care and Use Committees of the Massachusetts Institute of Technology.

#### Loss-of-function assay in human CD34<sup>+</sup> erythroid culture system

The lentiviral backbone vector pLKO.1 and packaging plasmids were transfected into 293T cells. Supernatants containing viral particles were harvested at 48 and 72 h. Primary human CD34<sup>+</sup> haematopoietic cells were transduced with lentivirus 1 day after thawing with the presence of 2  $\mu$ g ml<sup>-1</sup> polybrene (Sigma). Twenty four hours after viral transduction, cells were selected by growing in culture medium containing 1  $\mu$ g ml<sup>-1</sup> puromycin (Sigma) for 2 days.

For RPS19 knockdown,  $5 \times 10^5$  CD34 cells at day 1 of culture were infected with lentivirus encoding GFP and a shRNA targeting human RPS19 (ref. 21) or a scrambled shRNA. Cells were then treated with DMSO or 0.01  $\mu$ M, 0.1  $\mu$ M or 1  $\mu$ M GW7647. After 48 h, GFP-positive cells were isolated by flow cytometry and returned to culture. Colony-forming assays were conducted at day 6. CD71 expression was analysed at day 9 and CD235 expression was analysed at day 21 to determine the percentage of erythroid cells. Total cell numbers were also counted at the end of each differentiation stage.

**ChIP-seq and de novo motif discovery.** Mouse ChIP-seq experiments in BFU-Es were conducted as described before<sup>24</sup>.  $10^7$  mouse BFU-E cells purified from E14.5 fetal liver with or without treatment were used per immunoprecipitation. Besides specific antibodies, we included species-matched IgG as control, and species-matched IgG yielded little to no signals. Purified DNA was prepared for sequencing according to a modified version of the Solexa genomic DNA protocol. ChIP-seq fragments as well as inputs were barcoded and sequenced on an Illumina HiSeq sequencer. Approximately 30 million 40-nucleotide-long single reads per sample were obtained. Adapters were removed and reads shorter than 20 nucleotides were discarded. Reads were mapped with Bowtie1 (ref. 25). Peaks were called with MACS 1.4 using the corresponding input for each sample and “mfold” set to 5,30 (ref. 26). We selected peaks with fold enrichment  $\geq 10$ . Plots showing density of reads around the peak summits were done with ngsplo. For de novo motif discovery, Homer was used with default setting to find motifs in GR and PPAR- $\alpha$  overlapping peaks<sup>27</sup>.

**RNA-seq.** BFU-E cells from E14.5 mouse fetal livers were isolated as described<sup>3</sup>. Total RNA was purified from the mouse BFU-Es. Samples for paired-end mRNA-seq were prepared using the Solexa kit according to the manufacturer's instructions. RNA samples from two replicates of cells untreated, or treated with DEX  $\pm$  GW7647 for 12 h were sequenced on an Illumina HiSeq sequencer. Around 50 million 100-nucleotide paired-end reads were obtained from each sample. Reads were trimmed to remove low quality reads using FASTQ quality trimmer with a quality threshold of 20 ( $-t$  20) and a minimum length to keep 25 nucleotides ( $-l$  25). Reads that still had both pairs after the trimming step were mapped with TopHat<sup>28</sup> using gene models from ENSEMBL Genes 67, *Mus musculus* genes NCBI37 (*Mus\_musculus.NCBI37.67*). The number of reads mapped to each gene was obtained with HTseq-count. Differential expression was assayed using DE-seq<sup>29</sup>.

For each of the 719 peaks bound by both GR and PPAR- $\alpha$  in the presence of GW7647 and DEX, we found the closest gene using the closestBed tool<sup>30</sup>. We filtered out any gene at a distance higher than 10 kb from a peak.

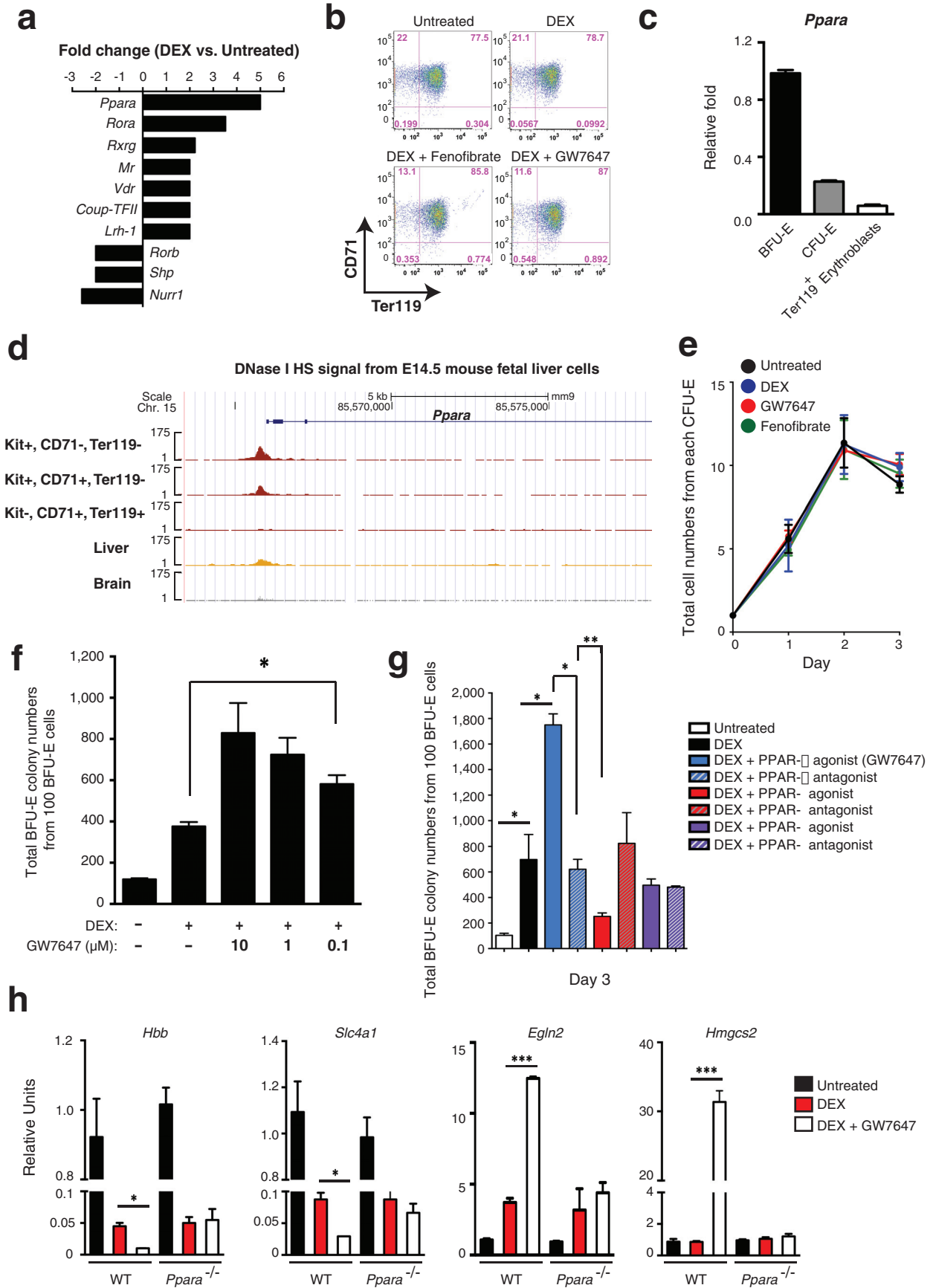
**Immunoprecipitation.** Mouse BFU-Es were isolated from E14.5 mouse fetal livers. Cells were cultured in SFELE medium alone or treated with 100 nM DEX with or without 10  $\mu$ M GW7647 for 12 h. Whole-cell lysates were prepared from  $4 \times 10^7$  cells using RIPA lysis buffer (sc-24948, Santa Cruz Biotechnology) with protease inhibitor cocktail. The lysates were pre-cleared by incubating with protein A-Sepharose for 1 h at 4 °C and centrifugation. The supernatant was immunoprecipitated with 1  $\mu$ g rabbit IgG or anti-GR antibody overnight at 4 °C. Immune complexes were collected by incubation with protein A-Sepharose for 4 h at 4 °C



and washed five times at 4 °C with lysis buffer. The immune complexes adsorbed to the beads were centrifuged and the supernatant was removed. Fifty microlitres of 1× loading buffer was added to the samples and boiled at 95 °C for 5 min. Proteins were resolved by SDS-PAGE and immunoblotted by GR and PPAR- $\alpha$  antibodies.

**Quantitative real-time RT-PCR.** Total RNA from mouse BFU-Es was purified with TRIzol (Invitrogen). cDNA was prepared from 1  $\mu$ g RNA. Reaction mixtures (15  $\mu$ l) contained 2.0  $\mu$ l of cDNA, 7.5  $\mu$ l of SYBR green master mix (Applied Biosystems) and appropriate primers. Product was monitored by SYBR green fluorescence. Control reactions lacking RT yielded little to no signal. Relative expression levels were determined from a  $\Delta\Delta C_t$  method and were normalized to 18S rRNA expression.

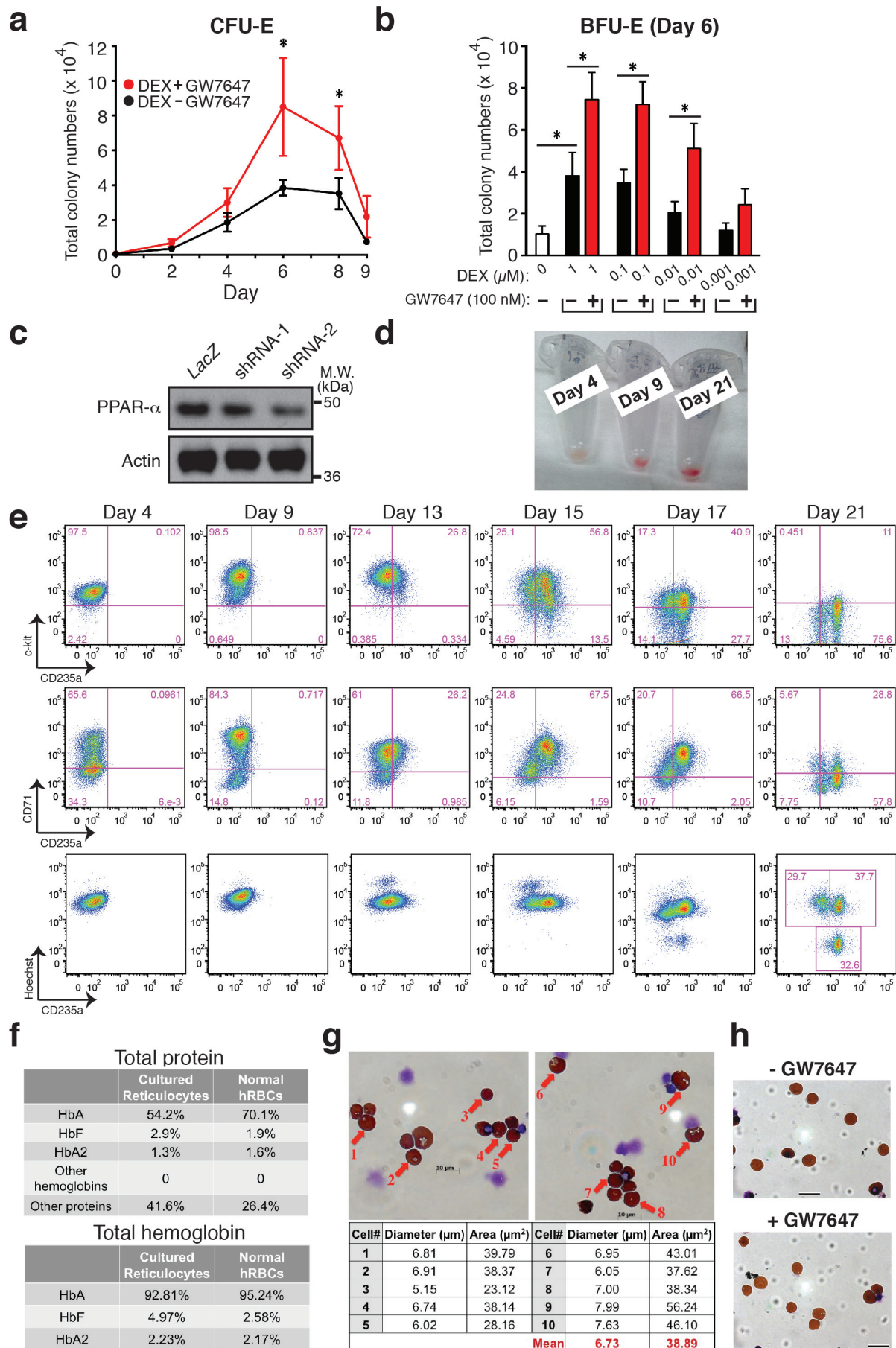
25. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
26. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
27. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
28. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
29. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
30. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).



**Extended Data Figure 1 | PPAR- $\alpha$  agonist GW7647 does not have adverse effects on erythroid differentiation and has no effects on CFU-E cells.**

**a**, Gene expression changes of nuclear receptors in BFU-E cells from RNA-seq results published previously<sup>10</sup>. **b**, Flow cytometry analyses of CD71 and Ter119 markers to demonstrate erythroid differentiation of mouse BFU-E cells after 9 days of culture with the indicated additions. **c**, *Ppara* gene expression in BFU-E, CFU-E and Ter119<sup>+</sup> erythroblasts. BFU-E, CFU-E and Ter119<sup>+</sup> erythroid cells were isolated from E14.5 mouse fetal livers as described<sup>24</sup>. Total RNA was purified for quantitative PCR analysis. *Ppara* gene expression was normalized to mouse 18S rRNA in different stages. Error bars represent mean  $\pm$  s.d. from three independent experiments. **d**, DNase I hypersensitivity (HS) analysis at *Ppara* promoter region in different mouse cells from Encode. **e**, Production of mouse erythroblasts from isolated CFU-E cells. Wild-type mouse CFU-E cells from E14.5 fetal livers were untreated (black line) or treated with DEX (blue line), GW7647 (red line) or fenofibrate (green line). Error bars represent

mean  $\pm$  s.d. from three independent experiments. **f**, Colony-forming assays were conducted at 48 h after compound treatment to determine BFU-E colony numbers from 100 mouse BFU-E cells cultured under the indicated conditions. \* $P < 0.05$ , Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments. **g**, At day 3, BFU-E colony numbers from 100 purified mouse BFU-E cells were quantified by colony forming assays. 100 purified mouse BFU-E cells were untreated or treated with DEX alone or DEX in combination with agonists or antagonists targeting PPAR receptors ( $\alpha$ ,  $\gamma$  or  $\beta$ ). BFU-E colonies were quantified after 8 days in culture. \* $P < 0.05$ , \*\* $P < 0.01$ , Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments. **h**, Real-time PCR analysis of gene expression in DEX-treated and DEX+GW7647-treated wild-type or *Ppara*<sup>-/-</sup> mouse BFU-E cells. \* $P < 0.05$ , \*\*\* $P < 0.001$ , Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments.

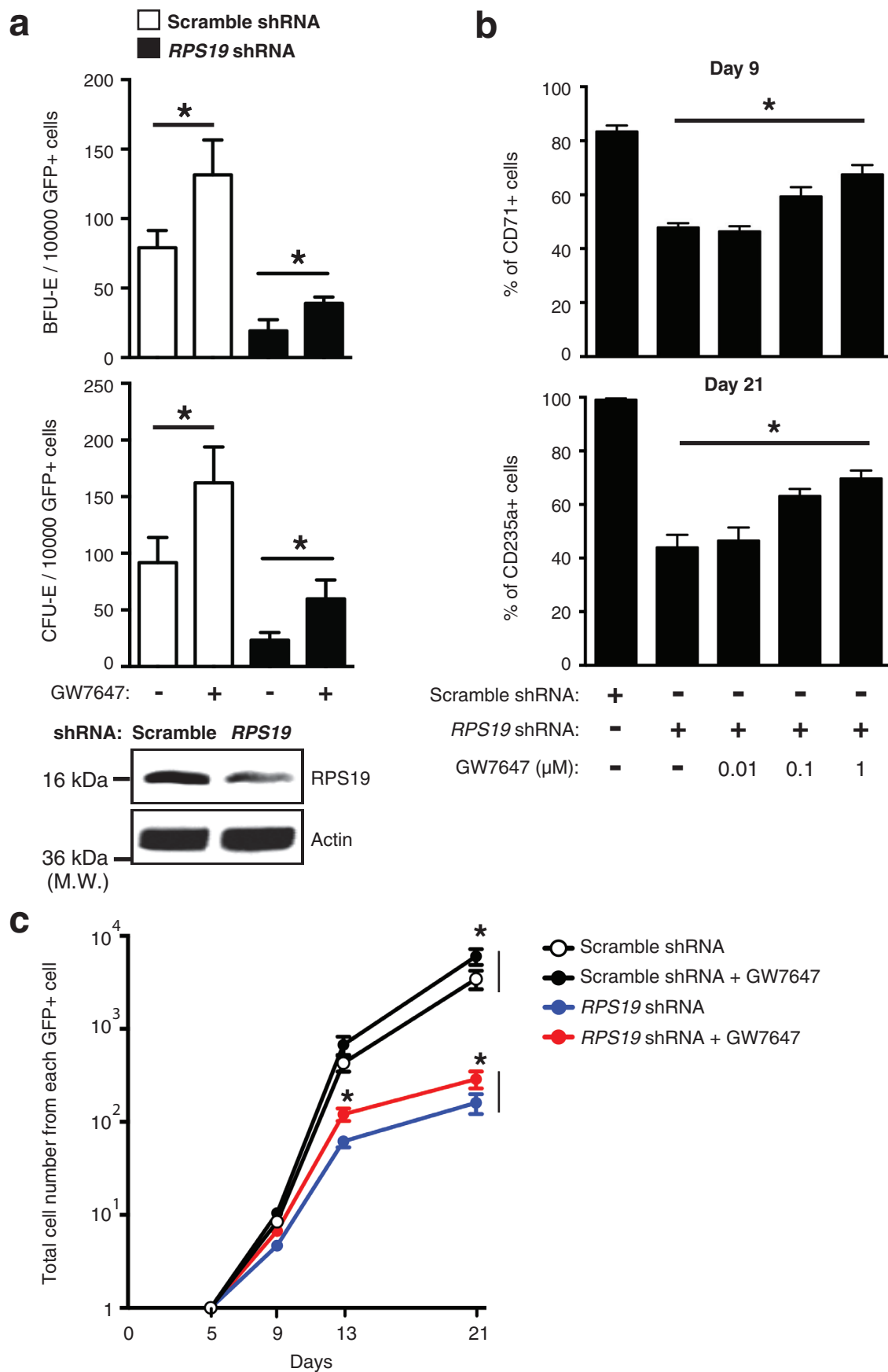




**Extended Data Figure 2 | Human CD34<sup>+</sup> erythroid differentiation system.**

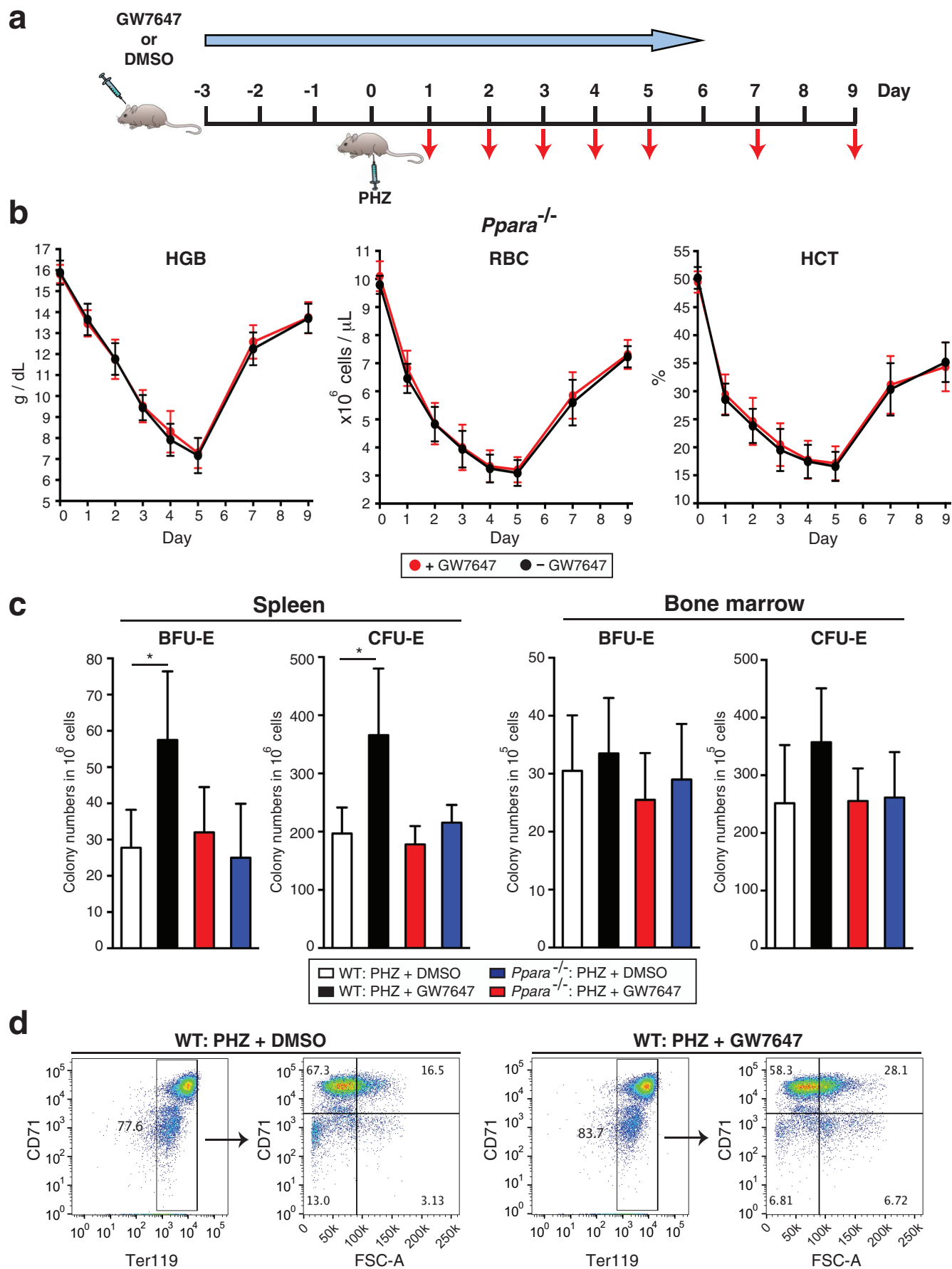
**a**, Total CFU-E colonies formed during days 0–9. CFU-E colony numbers were quantified by plating 1,000 cells from various time points during days 0–9 of the human CD34<sup>+</sup> erythroid culture on methylcellulose. CFU-E colonies were quantified after 12–14 days. Total CFU-E colony numbers in culture under conditions without GW7647 (black line) or with GW7647 (red line) were calculated using the total cell numbers at corresponding time points in Fig. 2a. **b**, Human CD34<sup>+</sup> cells were treated at day 1 with 100 nM GW7647 with or without DEX at the concentration indicated in the figure. At day 6, total cell numbers were counted and cells were collected for BFU-E colony assays. **c**, Protein expression of PPAR- $\alpha$  demonstrating shRNA knockdown efficiency via lentiviral transduction. *LacZ* shRNA is used as a control. shRNA-1 and -2 are both specific for PPAR- $\alpha$ . shRNA-2 has higher knockdown efficiency. **d**, Cell pellets of 1 million cells demonstrating haemoglobin accumulation during the differentiation process. **e**, Flow cytometry analyses of erythroid

markers during the 21-day human CD34<sup>+</sup> erythroid culture. Top, c-KIT versus CD235a; middle, CD71 versus CD235a. Note the sequential induction of c-kit, CD71 and CD235a, as well as the sequential downregulation of c-kit and CD71. Bottom, enucleated reticulocytes are CD235a<sup>+</sup>Hoechst<sup>-</sup>, nuclei are CD235a<sup>-</sup>Hoechst<sup>+</sup>, and nucleated erythroblasts are CD235a<sup>+</sup>Hoechst<sup>+</sup>. Enucleation rate is  $32.6/(32.6+37.7) \times 100\% = 46.4\%$ . **f**, Summary of high-performance liquid chromatography (HPLC) results using haemolysates of cultured reticulocytes and normal human RBCs (control). Top, total protein composition of haemolysates. Bottom, haemoglobin composition of haemolysates. Cultured reticulocytes contain more than 90% of adult globins. **g**, Size measurement of enucleated reticulocytes by both diameter and area. Scale bar, 10  $\mu\text{m}$ . **h**, Benzidine–Giemsa staining of human reticulocytes cultured with or without GW7647. Scale bar, 12  $\mu\text{m}$ . \* $P < 0.05$ , Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments.



**Extended Data Figure 3 | GW7647 increases erythroid progenitors and CD235a<sup>+</sup> cells in RPS19-knockdown human progenitor cells.** **a**, Human CD34<sup>+</sup> haematopoietic progenitors were transduced with lentivirus encoding GFP and either a scrambled shRNA or an shRNA targeting RPS19. Then transduced cells were treated with or without 100 nM GW7647. After 48 h, GFP<sup>+</sup> cells were sorted by FACS and plated for BFU-E and CFU-E colony-forming assays. RPS19 knockdown efficiency is shown at the bottom. M.W., molecular weight. \* $P < 0.05$ , Student's  $t$ -test. Error bars represent mean  $\pm$  s.d.

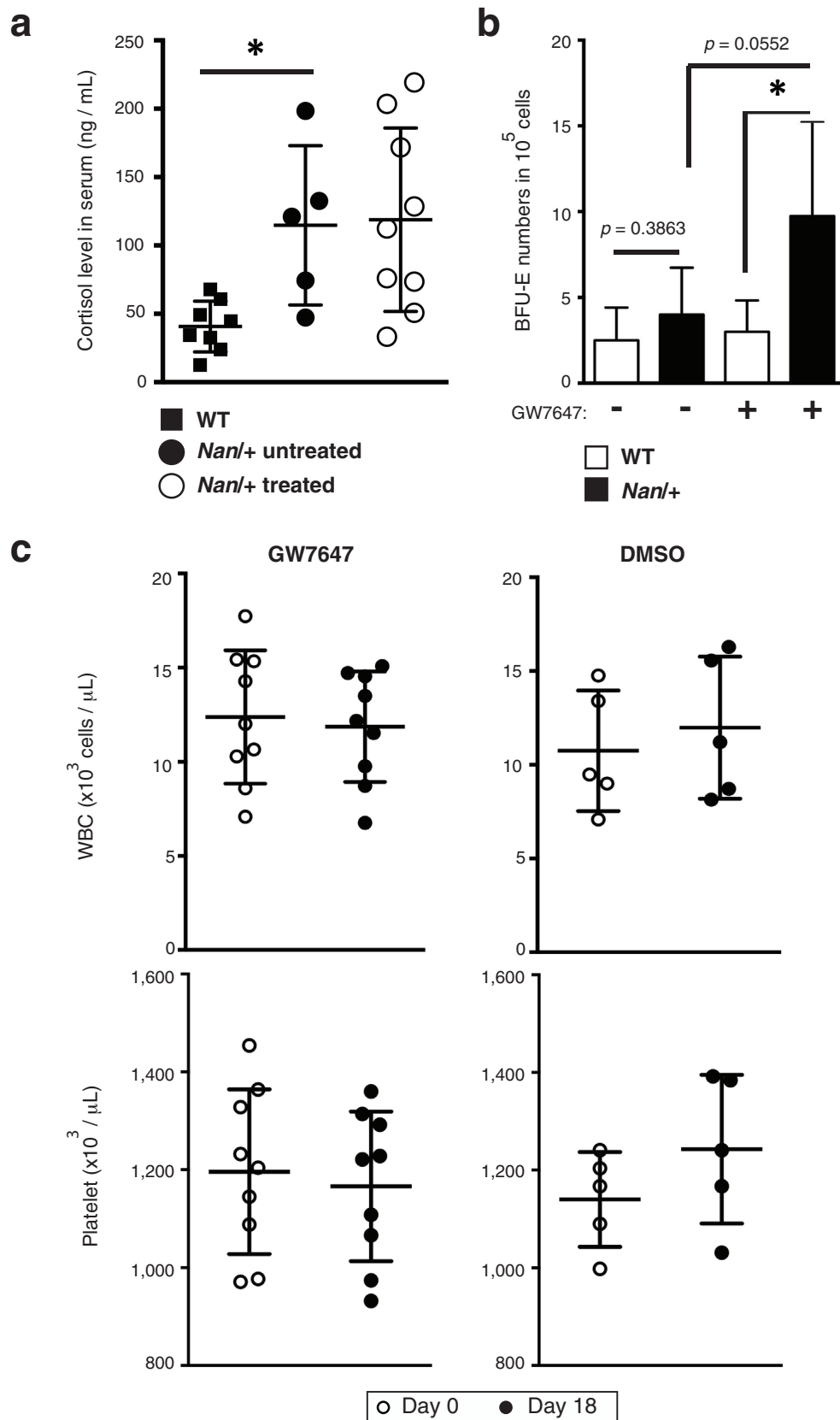
from three independent experiments. **b**, Sorted GFP<sup>+</sup> cells were returned to culture with the indicated concentration of GW7647. Top, percentage of CD71<sup>+</sup> cells at day 9 was determined by FACS. Bottom, percentage of CD235a<sup>+</sup> cells at day 21 was determined by FACS. \* $P < 0.05$ , Student's  $t$ -test. Error bars represent mean  $\pm$  s.d. from three independent experiments. **c**, Total cell numbers generated from one GFP-positive cell at the indicated times of culture. \* $P < 0.05$ , Student's  $t$ -test. Error bars represent mean  $\pm$  s.d. from three independent experiments.





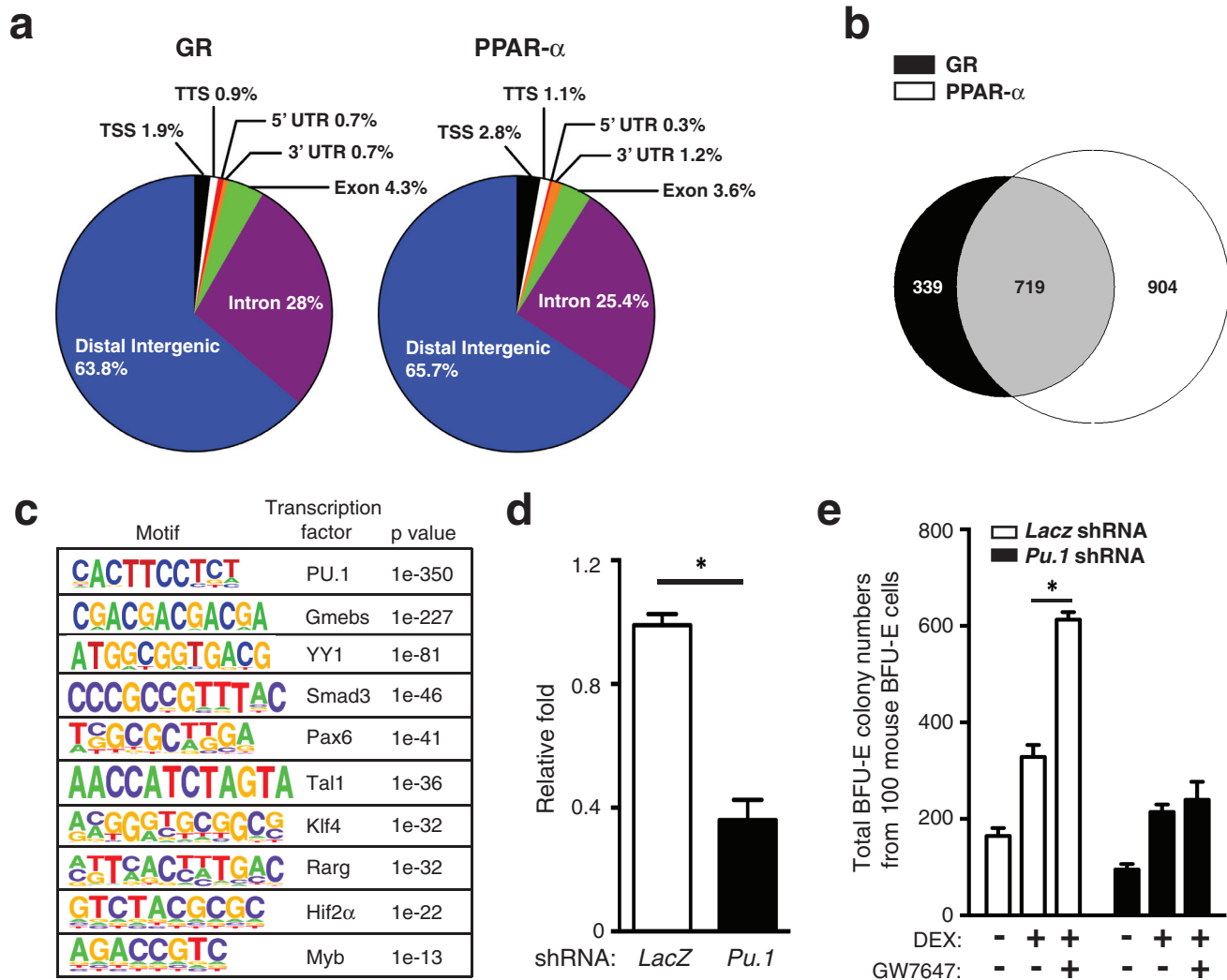
**Extended Data Figure 4 | GW7647 improves anaemia in two mouse models of anaemia.** **a**, Experimental scheme for PHZ- induced haemolytic anaemia, used also in Fig. 3a. Wild-type or *Ppara*<sup>-/-</sup> mice were pretreated with DMSO (control) or GW7647 (100 µg kg<sup>-1</sup>) for 3 days (days -3 to -1) before PHZ injection on day 0. Mice were subject to daily DMSO or GW7647 injections during days 0-6. Red arrows indicate days of blood sample collection. **b**, *Ppara*<sup>-/-</sup> mice were treated with DMSO or GW7647 and then injected with PHZ. Haemoglobin (HGB), red blood cell numbers (RBC) and haematocrit

(HCT) were measured on the days indicated. Error bars represent mean ± s.d. from six mice. **c**, BFU-E and CFU-E colony-forming assays of spleen or bone marrow cells. Wild-type or *Ppara*<sup>-/-</sup> mice were treated with PHZ and DMSO (control) or GW7647 (100 µg kg<sup>-1</sup>) as described in earlier. \**P* < 0.05, Student's *t*-test. Error bars represent mean ± s.d. from three independent experiments. **d**, Spleen and bone marrow cells were harvested. Representative flow cytometry analysis of spleen erythroblasts isolated from GW7647- or DMSO-treated wild-type mice at day 3 after PHZ injection. FSC-A, forward scatter area.



**Extended Data Figure 5 | GW7647 increases BFU-E numbers in *Nanl*<sup>+</sup> mutant mice.** **a**, Corticosteroid levels in serum were measured in wild-type and *Nanl*<sup>+</sup> mutant mice. Each dot represents one mouse. \* $P < 0.05$ , Student's *t*-test. Error bars represent mean  $\pm$  s.d. from all mice. **b**, Increase of BFU-E numbers in spleens from GW7647-treated wild-type and *Nanl*<sup>+</sup> mutant mice

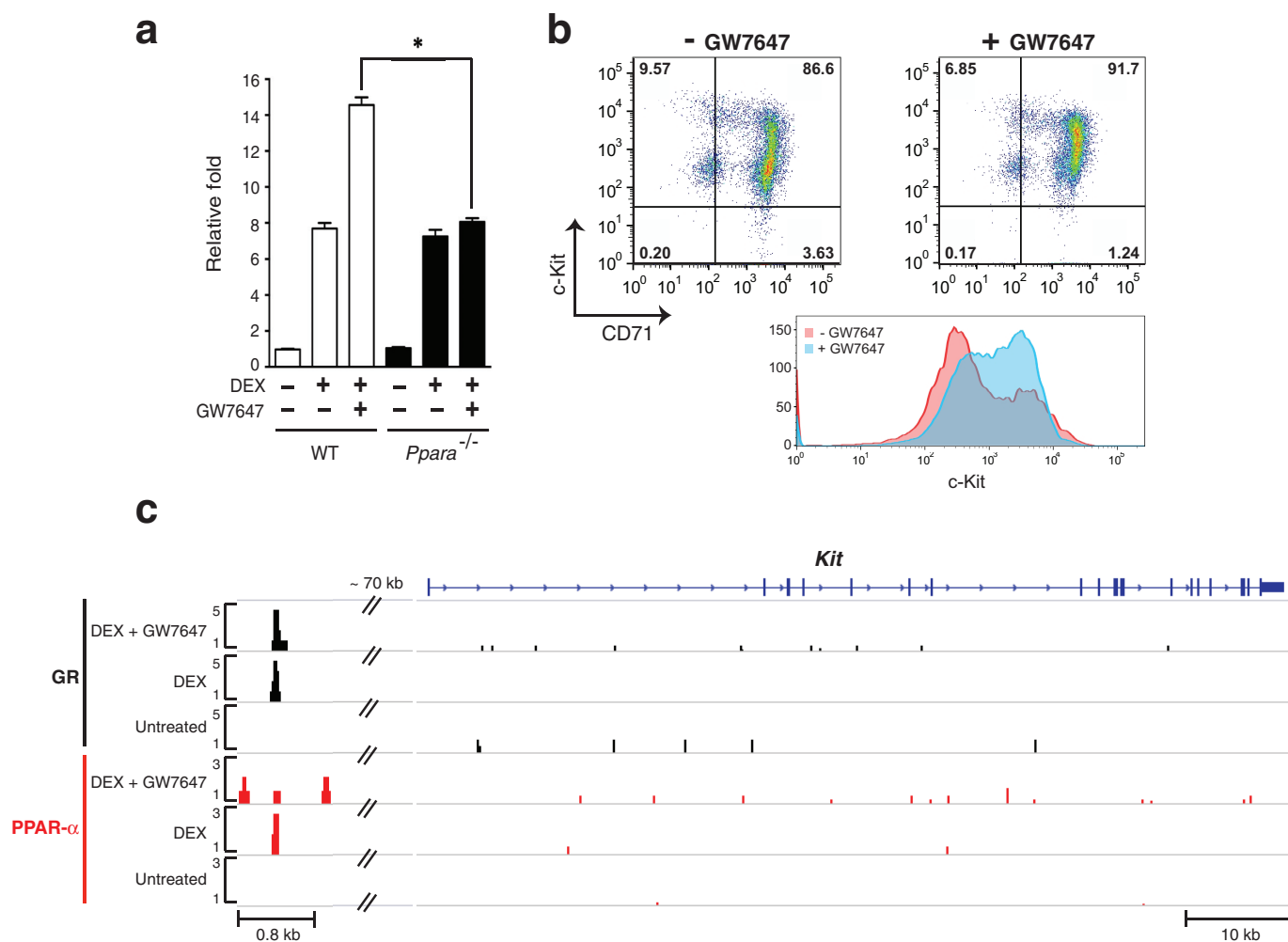
at day 18. \* $P < 0.05$ , Student's *t*-test. **c**, Total numbers of white blood cells (WBC) and platelets from peripheral blood samples were measured at day 0 and day 18. Each dot represents one mouse. Error bars represent mean  $\pm$  s.d. from all mice.



### Extended Data Figure 6 | Bioinformatic analyses of mouse BFU-E cells.

**a**, Genome-wide distribution of GR and PPAR- $\alpha$  chromatin occupancy sites in BFU-E cells. ChIP-seq analyses of GR and PPAR- $\alpha$  occupancy in mouse BFU-E cells isolated from DEX- and GW7647-treated wild-type E14.5 fetal livers. TSS, transcription start site; TTS, transcription termination site; UTR, untranslated region. Distal intergenic, regions greater than 3 kb from TSS. **b**, Venn diagram showing overlap between GR and PPAR- $\alpha$  chromatin occupancy sites. **c**, *De novo* motif searching of the overlapped chromatin sites occupied by GR and PPAR- $\alpha$ . The table depicts transcription-factor-binding motifs enriched at GR and PPAR- $\alpha$  overlapping sites relative to genomic background and associated *P*

values. **d**, Real-time PCR analysis of *Pu.1* gene expression in mouse BFU-E cells transduced with virus encoding either *LacZ* shRNA or *Pu.1* shRNA. \**P* < 0.05, Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments. **e**, Colony-forming assays were conducted to determine BFU-E colony numbers from 100 mouse BFU-E cells infected with virus encoding either *LacZ* shRNA or *Pu.1* shRNA. Cells were cultured in SFELE medium with or without DEX  $\pm$  GW7647. Colony-forming assays were performed at 48 h. \**P* < 0.05, Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments.



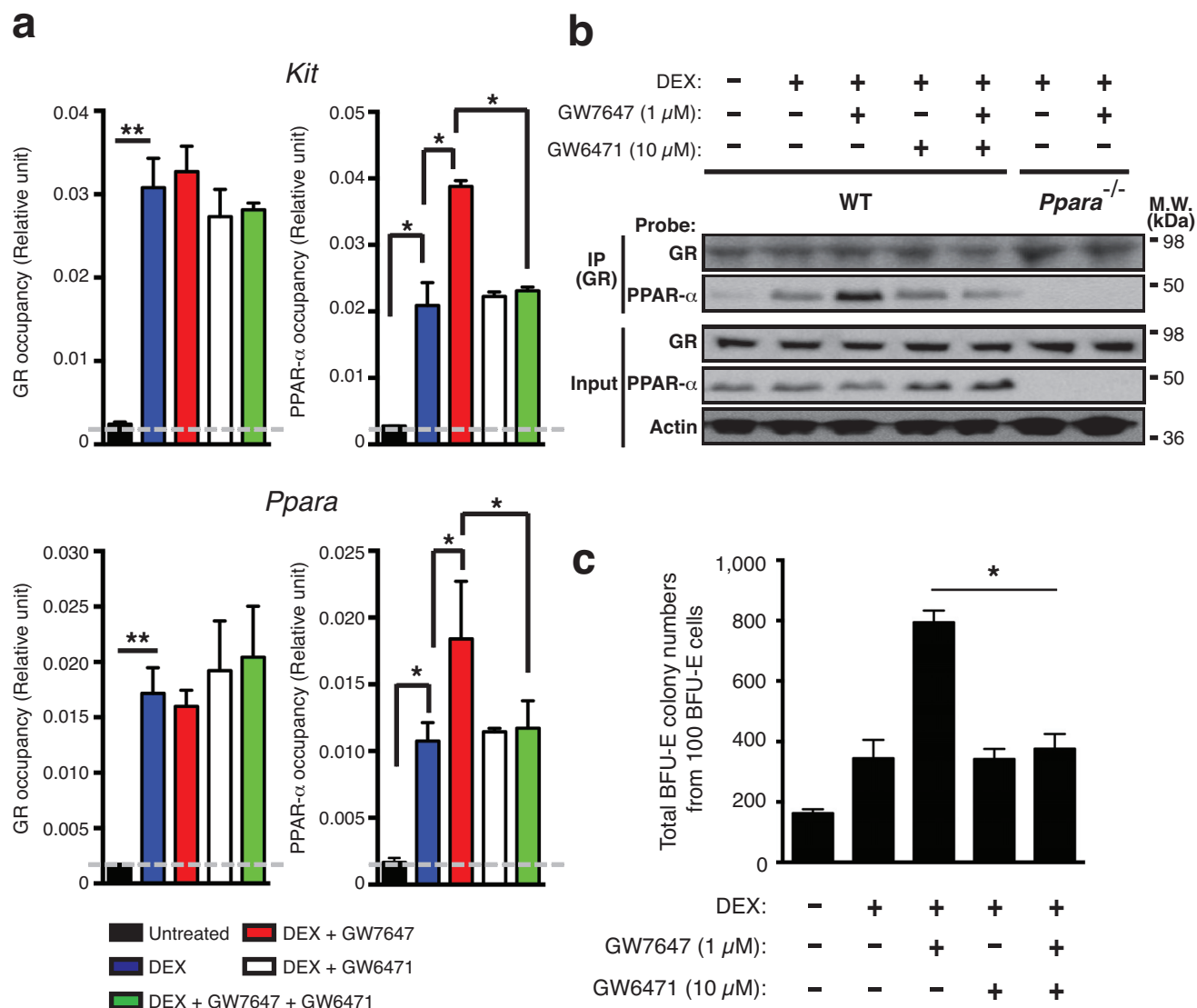
**Extended Data Figure 7 | *Kit* is a target gene of GR and PPAR- $\alpha$ .** **a**, Real-time RT-PCR analysis of *Kit* gene expression in wild-type and *Ppara*<sup>-/-</sup> mouse BFU-E cells untreated or treated with DEX with or without the addition of GW7647. \* $P < 0.05$ , Student's  $t$ -test. Error bars represent mean  $\pm$  s.d. from three independent experiments. **b**, Human CD34<sup>+</sup> cells were treated with or

without GW7647 as described in Fig. 2. Top, at day 9 of culture, cell-surface KIT and CD71 expression were analysed by flow cytometry. Bottom, a representative histogram of KIT expression in cells treated or untreated with GW7647. **c**, ChIP-seq occupancy signal map of GR and PPAR- $\alpha$  across the *Kit* locus in BFU-E cells.



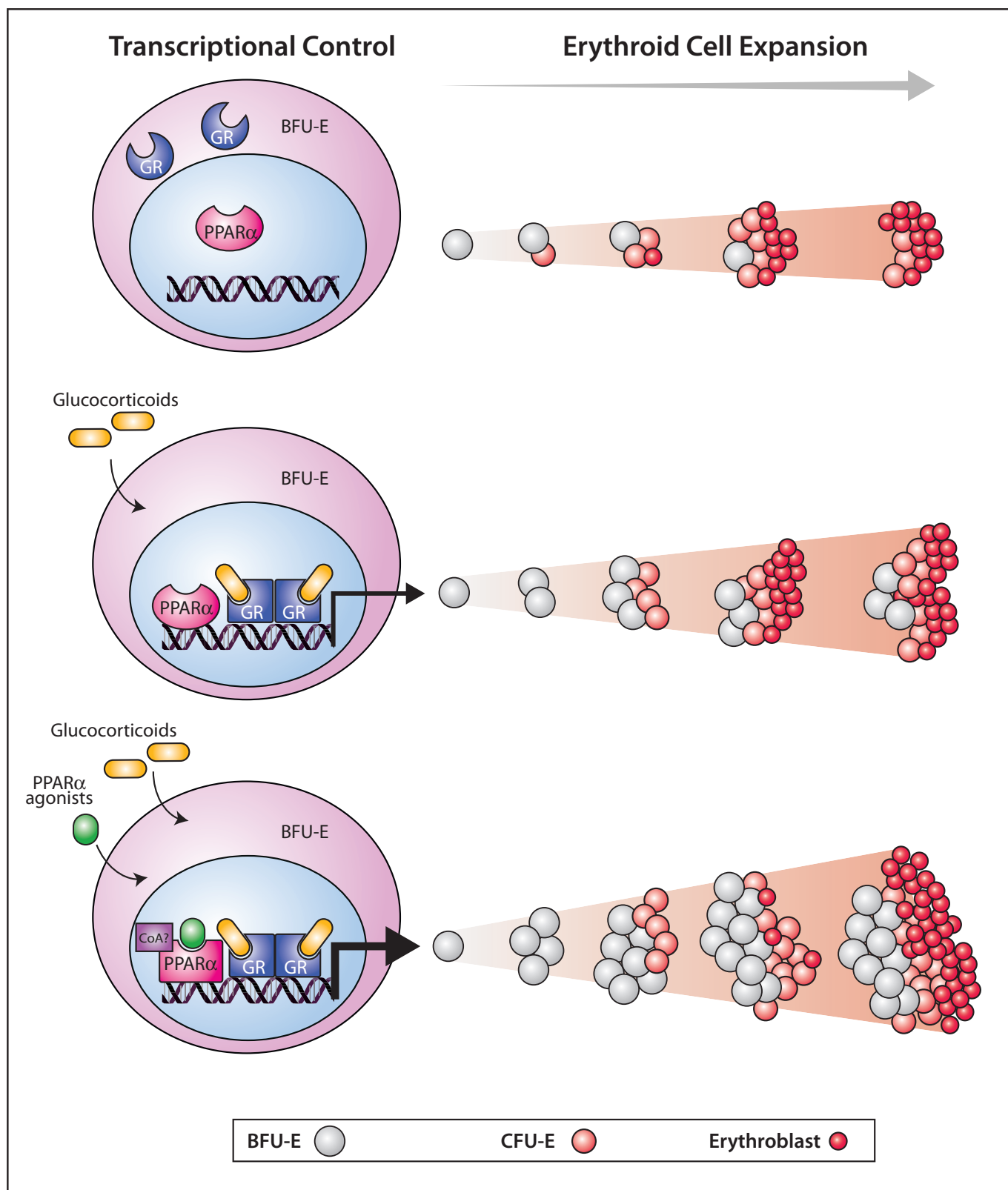
| KEGG pathways                          | P value              |
|--|----------------------|
| Hedgehog signaling pathway             | 2.6 e <sup>-8</sup>  |
| Calcium signaling pathway              | 7.19 e <sup>-6</sup> |
| Cytokine-cytokine receptor interaction | 1.28 e <sup>-4</sup> |
| Cell adhesion molecules                | 9.11 e <sup>-4</sup> |
| Hematopoietic cell lineage             | 1.46 e <sup>-3</sup> |
| Wnt signaling pathway                  | 1.55 e <sup>-3</sup> |

**Extended Data Figure 8 | Pathways modulated by PPAR- $\alpha$  activation.** Pathway analysis of RNA-seq data of genes that are up- or downregulated by more than 50%, comparing cultures treated with DEX alone or DEX+GW7647.



**Extended Data Figure 9 | PPAR- $\alpha$  antagonist interferes with the function of PPAR- $\alpha$  but not GR.** **a**, Quantitative ChIP analysis of GR and PPAR- $\alpha$  occupancy at *Kit* and *Ppara* loci in mouse BFU-E cells following the indicated treatments. Units are arbitrary; signals using rabbit IgG are represented by grey dot lines across the plots. \* $P < 0.05$ , \*\* $P < 0.01$ , Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments. **b**, Co-immunoprecipitation (IP) measuring interaction between GR and PPAR- $\alpha$  in mouse BFU-E cells isolated from E14.5 fetal livers in wild-type or

*Ppara*<sup>-/-</sup> mice. BFU-E cells were untreated, or treated with DEX with or without GW7647 with or without GW6471. Whole-cell lysates were incubated with anti-GR antibody and immunoprecipitates were probed with specific antibodies as indicated. M.W., molecular weight. **c**, Colony-forming assays to determine BFU-E colony numbers from 100 mouse BFU-E cells cultured with the indicated treatments. \* $P < 0.05$ , Student's *t*-test. Error bars represent mean  $\pm$  s.d. from three independent experiments.



**Extended Data Figure 10 | Model of synergism between PPAR- $\alpha$  and GR to promote BFU-E self-renewal.** BFU-E cells normally undergo limited self-renewal to generate CFU-Es, which can differentiate into erythroblasts. GR is sequestered in the cytoplasm without glucocorticoids such as DEX. Upon glucocorticoid treatment, liganded GR will be translocated into the nucleus and bind to chromatin to regulate gene transcription important for BFU-E self-

renewal. PPAR- $\alpha$  is often recruited to chromatin sites that are in close proximity to GR by glucocorticoid treatment alone without any function on BFU-E self-renewal. Upon glucocorticoid and PPAR- $\alpha$  agonist co-treatment, activated PPAR- $\alpha$  interacts with GR to modulate GR transcriptional activity. This leads to enhanced BFU-E self-renewal, and over time generates more CFU-Es and erythroblasts.

# A naturally occurring variant of the human prion protein completely prevents prion disease

Emmanuel A. Asante<sup>1</sup>, Michelle Smidak<sup>1</sup>, Andrew Grimshaw<sup>1</sup>, Richard Houghton<sup>1†</sup>, Andrew Tomlinson<sup>1</sup>, Asif Jeelani<sup>1</sup>, Tatiana Jakubcova<sup>1</sup>, Shyma Hamdan<sup>1</sup>, Angela Richard-Londt<sup>1</sup>, Jacqueline M. Linehan<sup>1</sup>, Sebastian Brandner<sup>1</sup>, Michael Alpers<sup>1,2</sup>, Jerome Whitfield<sup>1,2</sup>, Simon Mead<sup>1</sup>, Jonathan D. F. Wadsworth<sup>1</sup> & John Collinge<sup>1</sup>

Mammalian prions, transmissible agents causing lethal neurodegenerative diseases, are composed of assemblies of misfolded cellular prion protein (PrP)<sup>1</sup>. A novel PrP variant, G127V, was under positive evolutionary selection during the epidemic of kuru—an acquired prion disease epidemic of the Fore population in Papua New Guinea—and appeared to provide strong protection against disease in the heterozygous state<sup>2</sup>. Here we have investigated the protective role of this variant and its interaction with the common, worldwide M129V PrP polymorphism. V127 was seen exclusively on a M129 PRNP allele. We demonstrate that transgenic mice expressing both variant and wild-type human PrP are completely resistant to both kuru and classical Creutzfeldt–Jakob disease (CJD) prions (which are closely similar) but can be infected with variant CJD prions, a human prion strain resulting from exposure to bovine spongiform encephalopathy prions to which the Fore were not exposed. Notably, mice expressing only PrP V127 were completely resistant to all prion strains, demonstrating a different molecular mechanism to M129V, which provides its relative protection against classical CJD and kuru in the heterozygous state. Indeed, this single amino acid substitution (G→V) at a residue invariant in vertebrate evolution is as protective as deletion of the protein. Further study in transgenic mice expressing different ratios of variant and wild-type PrP indicates that not only is PrP V127 completely refractory to prion conversion but acts as a potent dose-dependent inhibitor of wild-type prion propagation.

Prions cause fatal neurodegenerative conditions such as scrapie in sheep, bovine spongiform encephalopathy (BSE) in cattle and CJD in humans<sup>1</sup>. The fundamental molecular process—seeded propagation of assemblies of misfolded host protein—is increasingly recognized as being of importance in all the major human neurodegenerative diseases<sup>3</sup>. There is a common polymorphism, present worldwide, in the substrate protein in human prion disease, human prion protein (PrP), where either methionine (M) or valine (V) is present at residue 129. MV heterozygosity provides relative protection against acquired, sporadic and some inherited prion diseases<sup>4–6</sup> and may have been selected during the evolution of modern humans by ancestral prion disease epidemics<sup>7</sup>. This protective effect is thought to relate to inhibition of homotypic protein–protein interactions<sup>5,8</sup>, although residue 129 also influences the propagation of particular prion strains via conformational selection<sup>9,10</sup>. Heterozygosity at another polymorphism, E219K, also provides resistance to CJD in Japan<sup>11</sup>.

Kuru was a devastating epidemic prion disease transmitted by endocannibalism and restricted to a remote area in Papua New Guinea. We reported a novel PrP variant (G127V) among unaffected individuals that appeared to be a resistance factor selected by the epidemic and unique to this region<sup>2</sup>. Given the proximity to residue 129, we considered that it may have a similar action to M129V; that is, blocking homotypic interactions and exerting its protective effect only in the heterozygous state. No PRNP codon 127VV homozygotes were

identified in the kuru-exposed population and V127 was always seen on an M129 allele<sup>2</sup>. We therefore generated multiple lines of transgenic mice, expressing only human PrP (HuPrP) on a congenic FVB/N PrP null (*Prnp*<sup>0/0</sup>) background, to investigate whether G127V was indeed protective and whether this protection was dependent on heterozygosity or was an intrinsic property of the variant protein. Furthermore, we investigated its interaction with the residue 129 polymorphism.

We generated two lines of transgenic mice homozygous for HuPrP V127: Tg(HuPrP V127M129/V127M129 *Prnp*<sup>0/0</sup>)-183 (hereafter referred to as V127M129 Tg183) and Tg(HuPrP V127M129/V127M129 *Prnp*<sup>0/0</sup>)-190 (V127M129 Tg190). PrP expression levels in homozygotes as compared to pooled normal human brain were twofold for V127M129 Tg183 and onefold for V127M129 Tg190. G127V heterozygous mice were derived by crossing these lines with FVB-congenic versions of Tg35 mice homozygous for HuPrP G127M129 (refs 12–15) or Tg152 mice homozygous for HuPrP G127V129 (refs 15–19), designated Tg(HuPrP G127M129/G127M129 *Prnp*<sup>0/0</sup>)-35c (G127M129 Tg35c) and Tg(HuPrP G127V129/G127V129 *Prnp*<sup>0/0</sup>)-152c (G127V129 Tg152c), respectively. G127M129 Tg35c and G127V129 Tg152c mice express wild-type HuPrP at two and six times, respectively, as compared to pooled normal human brain (Table 1). Extended Data Fig. 1 shows relative PrP<sup>C</sup> expression levels in all transgenic mice used in this study.

All transgenic lines were challenged by intracerebral prion inoculation from well-characterized and previously transmitted human prion disease cases including all three PRNP codon 129 genotypes and comprising four cases of kuru, twelve cases of classical CJD and two cases of variant CJD (vCJD).

Heterozygous HuPrP G127M129/V127M129 mice (Tg35c × Tg183 and Tg35c × Tg190) (Table 1), which have the genotype associated with disease resistance in the kuru-exposed human population, were completely resistant to all four kuru isolates (which included all three PRNP codon 129 genotypes and two molecular strain types) (Fig. 1a and Table 2a), while mice expressing wild-type HuPrP G127M129/G127M129 (Tg35c) (Fig. 1a and Table 2a) or G127V129/G127V129 (Tg152c) (Table 2b) were fully susceptible with 100% attack rates. This is consistent with the population genetic data suggesting kuru resistance of G127V individuals. We have previously reported that prion strains seen in kuru brain are indistinguishable from those seen in classical CJD patients<sup>19</sup>. Similarly, none of the four classical CJD isolates from patients of the PRNP genotype 129MM transmitted to either line of G127M129/V127M129 mice (Table 2a and Fig. 1b), while all four transmitted uniformly to Tg35c mice expressing wild-type HuPrP (Fig. 1b and Table 2a). Remarkably, however, some (~10%) G127M129/V127M129 mice developed clinical disease, and a larger number (~70%) showed evidence of subclinical infection (positive PrP immunohistochemistry and/or western blot for PrP<sup>Sc</sup>), on challenge with vCJD prions (Table 2d and Extended Data Fig. 2). vCJD is a novel BSE-derived prion strain<sup>17,20,21</sup>

<sup>1</sup>MRC Prion Unit, Department of Neurodegenerative Disease, UCL Institute of Neurology, London WC1N 3BG, UK. <sup>2</sup>Papua New Guinea Institute of Medical Research, Goroka, Eastern Highlands Province, Papua New Guinea. <sup>†</sup>Present address: Strangeways Research Laboratory, Worts' Causeway, Cambridge CB8 8UX, UK.



**Table 1 | Transgenic mice expressing G127 and V127 human PrP**

| Transgenic mice               | PRNP codon 127-129 genotype | Total HuPrP*<br>expression level (×) | WT† G127 allele<br>expression level (×) | Variant‡ V127 allele<br>expression level (×) | PrP ratio<br>G127:V127 |
|-------------------------------|-----------------------------|--------------------------------------|---|--|------------------------|
| <b>Parental lines</b>         |                             |                                      |   |  |                        |
| Tg35c                         | G127M129/G127M129           | 2                                    | 2                                       | –  | –                      |
| Tg152c                        | G127V129/G127V129           | 6                                    | 6                                       | –  | –                      |
| Tg183                         | V127M129/V127M129           | 2                                    | –                                       | 2  | –                      |
| Tg190                         | V127M129/V127M129           | 1                                    | –                                       | 1  | –                      |
| <b>F1 crosses</b>             |                             |                                      |   |  |                        |
| Tg35c × Tg183                 | G127M129/V127M129           | 2                                    | 1                                       | 1  | 1:1                    |
| Tg35c × Tg190                 | G127M129/V127M129           | 1.5                                  | 1                                       | 0.5  | 2:1                    |
| Tg152c × Tg183                | G127V129/V127M129           | 4                                    | 3                                       | 1  | 3:1                    |
| Tg152c × Tg190                | G127V129/V127M129           | 3.5                                  | 3                                       | 0.5  | 6:1                    |
| Tg152c <sup>+/-</sup> (Hemi)§ | G127V129                    | 3                                    | 3                                       | –  | –                      |

WT, wild type.

\* PrP expression level is relative to pooled 10% (w/v) normal human brain homogenate.

† Wild-type alleles are either G127M129 or G127V129.

‡ Variant allele is V127M129.

§ Generated by crossing Tg152c with FVB/PrP-null mice.

to which the population in the kuru-affected area of Papua New Guinea were not exposed.

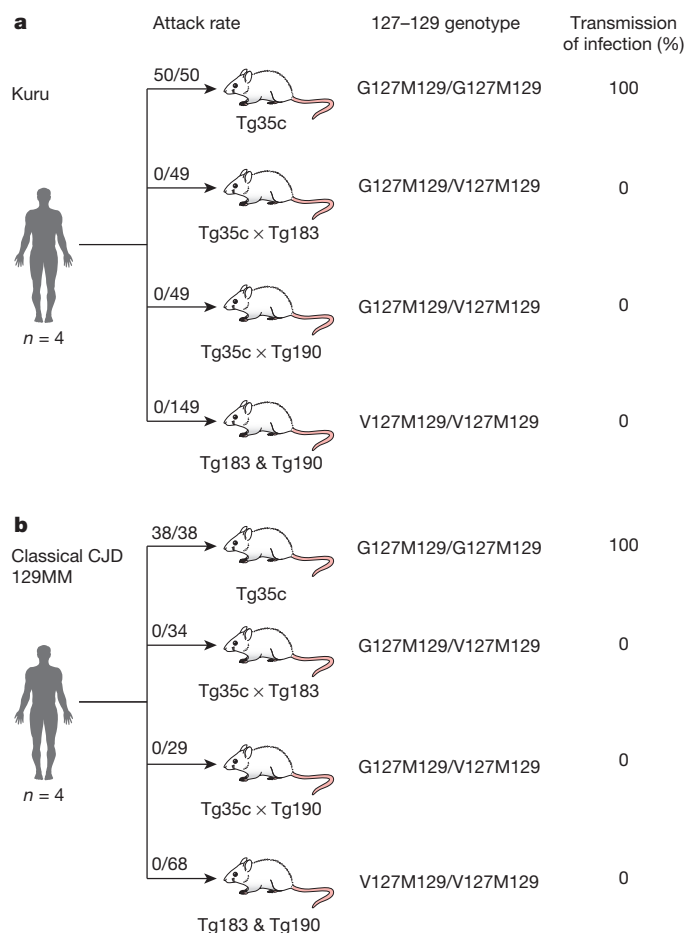
The key question with respect to whether the protective effect of V127 is conferred via a similar mechanism to the M129V polymorphism was addressed by challenge of mice homozygous for HuPrP

V127M129. While mice homozygous for wild-type HuPrP G127M129 or G127V129 are both highly susceptible to CJD prions, mice homozygous for HuPrP V127M129 were completely resistant to all 18 human prion disease isolates, including vCJD prions, with no clinical transmissions or evidence of subclinical infection (Fig. 1a, b, Table 3 and Extended Data Fig. 3).

We also challenged mice heterozygous at both residues 127 and 129 of HuPrP (G127V129/V127M129) produced by crossing Tg183 or Tg190 mice with Tg152c mice (Tables 1 and 2b). The higher level of expression of HuPrP V129 in Tg152c mice as compared to the other lines meant that there was a marked difference in the ratio of expression of the variant (V127M129) and wild-type (G127V129) HuPrP in these crosses (Table 1). While Tg183 and Tg35c had closely similar levels of expression of V127M129 and G127M129 HuPrP, respectively, such that the Tg35c × Tg183 cross closely modelled the human PRNP G127V genotype, the expression ratios of wild-type G127V129 to variant V127M129 HuPrP were approximately 6:1 and 3:1 for the Tg152c × Tg190 and Tg152c × Tg183 crosses, respectively (Table 1 and Extended Data Figs 1 and 4). Interestingly, in these crosses some transmissions of kuru and classical CJD prions were seen (Table 2b and Extended Data Figs 4 and 5). In the Tg152c × Tg183 crosses, with a threefold excess of wild-type PrP, subclinical infections were seen with four out of the twelve isolates used, and one kuru isolate (I10336) was associated with clinical disease (Table 2b). With the Tg152c × Tg190 crosses, however, showing a sixfold excess of wild-type PrP, transmissions were seen from all isolates, with the majority resulting in clinical disease (Table 2b).

To investigate further the effect of HuPrP V127 on prion propagation, we compared hemizygous Tg(HuPrP G127V129 *Prnp*<sup>0/0</sup>)-152c mice with the Tg152c × Tg183 and Tg152c × Tg190 crosses (Table 2c). The hemizygous line was challenged with two classical CJD and one kuru isolate that had resulted in a 100% attack rate in both crosses. Again, 100% attack rates were seen for all three isolates but the incubation periods were significantly shorter ( $P < 0.0001$ ; Table 2c) in the absence of HuPrP V127 than in either cross. These data, together with comparison of the two crosses themselves (Table 2b), are consistent with a dominant-negative effect of HuPrP V127 expression. Although dominant-negative inhibition has been reported for other natural polymorphic residues of mammalian PrP in transgenic mice<sup>22–24</sup>, none has shown complete prevention of prion conversion when expressed at a 1:1 ratio with wild-type PrP. Dominant-negative effects on yeast prion propagation have also been reported<sup>25</sup>.

Our transgenic modelling of the HuPrP G127V polymorphism demonstrates that it confers strong protection against prion disease in the heterozygous state. However, most importantly, the molecular basis of this effect is clearly distinct from that proposed for the well-established HuPrP M129V polymorphism, which is protective against developing sporadic CJD only in the heterozygous state: inhibition of homotypic protein–protein interactions during the process of prion



**Figure 1 | Transmission rates of human prions to transgenic mice homozygous or heterozygous for human PrP V127.** **a, b,** Transgenic mice were intracerebrally inoculated with brain homogenate from patients with kuru (**a**) or classical CJD (**b**). Codon 127 and 129 PRNP genotypes of the recipient mice are shown. G, glycine; M, methionine; V, valine. G127M129 is a wild-type human allele, and the V127M129 allele is seen only in humans from the kuru-exposed population of Papua New Guinea. Attack rate reports the total of clinically affected and subclinically infected mice as a proportion of the number of inoculated mice after prolonged (>600 days) post-inoculation periods. Primary prion transmission data are reported in Tables 2 and 3.

**Table 2 | Transmission of human prions to transgenic mice heterozygous for human PrP V127**

| Inoculum  |             |                               | Transmission data                |  |                                  |  |                                       |  |
|-----------|-------------|-------------------------------|----------------------------------|--|----------------------------------|--|---------------------------------------|--|
| Aetiology | Source code | Human PrP <sup>Sc</sup> type* | Attack rate†                     | Incubation period (days ± s.e.m.) or (days p.i.) | Attack rate†                     | Incubation period (days ± s.e.m.) or (days p.i.) | Attack rate†                          | Incubation period (days ± s.e.m.) or (days p.i.) |
| <b>a</b>  |             |                               | Tg35c G127M129/G127M129          |  | Tg35c x Tg183 G127M129/V127M129  |  | Tg35c x Tg190 G127M129/V127M129       |  |
| Kuru      | I516        | T3 VV                         | 14/14                            | 509 ± 56 (3)                                     | 0/11                             | >518–605   | 0/13                                  | >517–605   |
| Kuru      | I520        | T3 VV                         | 12/12                            | 558 ± 4 (4)                                      | 0/14                             | >525–622   | 0/14                                  | >460–622   |
| Kuru      | I10336      | T3 MV                         | 13/13                            | 454 ± 14 (10)                                    | 0/14                             | >522–606   | 0/11                                  | >504–607   |
| Kuru      | I518        | T2 MM                         | 11/11                            | 493 ± 32 (7)                                     | 0/12                             | >545–620   | 0/13                                  | >544–603   |
| iCJD (GH) | I035        | T1 MM                         | 10/10                            | 221 ± 3 (10)                                     | 0/10                             | >462–603   | 0/8                                   | >506–602   |
| sCJD      | I11058      | T1 MM                         | 8/8                              | 231 ± 3 (8)                                      | 0/8                              | >467–605   | 0/7                                   | >487–605   |
| iCJD (DM) | I026        | T2 MM                         | 10/10                            | 256 ± 9 (8)                                      | 0/8                              | >507–601   | 0/7                                   | >520–602   |
| sCJD      | I7040       | T2 MM                         | 10/10                            | 233 ± 3 (10)                                     | 0/8                              | >582–604   | 0/7                                   | >600–604   |
| <b>b</b>  |             |                               | Tg152c G127V129/G127V129         |  | Tg152c x Tg183 G127V129/V127M129 |  | Tg152c x Tg190 G127V129/V127M129      |  |
| Kuru      | I516        | T3 VV                         | 9/9                              | 218 ± 1 (6)                                      | 0/13                             | >447–608   | 9/15                                  | 543, 592   |
| Kuru      | I520        | T3 VV                         | 9/9                              | 196 ± 7 (7)                                      | 0/18                             | >491–616   | 13/13                                 | 498 ± 17 (5)                                     |
| Kuru      | I10336      | T3 MV                         | 15/15                            | 212 ± 3 (11)                                     | 15/15                            | 456 ± 3 (15)                                     | 15/15                                 | 316 ± 4 (14)                                     |
| Kuru      | I518        | T2 MM                         | 11/11                            | 211 ± 4 (8)                                      | 3/11                             | >494–620   | 13/14                                 | 468 ± 13 (9)                                     |
| sCJD      | I280        | T2 VV                         | 8/8                              | 203 ± 5 (4)                                      | 0/4                              | >600–602   | 10/10                                 | 559 ± 14 (8)                                     |
| sCJD      | I278        | T2 VV                         | 6/6                              | 236 ± 8 (6)                                      | 0/8                              | >424–602   | 8/9                                   | 582  |
| sCJD      | I284        | T2 MV                         | 9/9                              | 338 ± 4 (8)                                      | 0/9                              | >410–602   | 5/7                                   | >545–608   |
| sCJD      | I1478       | T2 MV                         | 10/10                            | 248 ± 8 (8)                                      | 0/6                              | >451–608   | 2/7                                   | >482–603   |
| sCJD      | I7394       | T3 VV                         | 9/9                              | 213 ± 1 (8)                                      | 7/7                              | >537–677   | 10/10                                 | 427 ± 8 (10)                                     |
| sCJD      | I764        | T3 MV                         | 10/10                            | 219 ± 5 (7)                                      | 9/10                             | >368–588   | 10/10                                 | 461 ± 7 (9)                                      |
| iCJD (GH) | I2651       | T3 VV                         | 8/8                              | 200 ± 3 (6)                                      | 0/6                              | >567–607   | 7/7                                   | 538 ± 2 (5)                                      |
| iCJD (GH) | I020        | T3 MV                         | 9/9                              | 211 ± 2 (6)                                      | 7/7                              | >489–602   | 10/10                                 | 399 ± 8 (9)                                      |
| <b>c</b>  |             |                               | Tg152c x Tg183 G127V129/V127M129 |  | Tg152c x Tg190 G127V129/V127M129 |  | Tg152c <sup>+/-</sup> (Hemi) G127V129 |  |
| iCJD (GH) | I020        | T3 MV                         | 7/7                              | >489–602   | 10/10                            | 399 ± 8‡ (9)                                     | 14/14                                 | 252 ± 3‡ (14)                                    |
| sCJD      | I7394       | T3 VV                         | 7/7                              | >537–677   | 10/10                            | 427 ± 8‡ (10)                                    | 12/12                                 | 245 ± 5‡ (12)                                    |
| Kuru      | I10336      | T3 MV                         | 15/15                            | 456 ± 3 (15)                                     | 15/15                            | 316 ± 4‡ (14)                                    | 15/15                                 | 222 ± 2‡ (15)                                    |
| <b>d</b>  |             |                               | Tg35c G127M129/G127M129          |  | Tg35c x Tg183 G127M129/V127M129  |  | Tg35c x Tg190 G127M129/V127M129       |  |
| vCJD      | I342        | T4 MM                         | 15/15                            | >426–603   | 9/15                             | 467, 556   | 17/18                                 | 419 ± 17 (3)                                     |
| vCJD      | I7042       | T4 MM                         | 14/14                            | 559, 561   | 4/13                             | >496–607   | 10/12                                 | 596  |

iCJD, iatrogenic CJD; sCJD, sporadic CJD; GH, growth hormone; DM, dura mater; p.i., post inoculation.

\* According to classification of ref. 28.

† Attack rate is defined as the total number of both clinically affected and subclinically infected mice as a proportion of the total number of inoculated mice. Subclinical prion infection was assessed by immunoblotting and/or immunohistochemical examination of brain. Incubation periods are reported for clinically affected mice in days; where  $n \geq 3$  the mean ± s.e.m. is reported with the number of mice contributing to the mean shown in parentheses, otherwise individual incubation times are given. In groups where no clinical transmission of prion disease was observed, the attack rate represents subclinical infection only and the interval between inoculation and death (from either senescence, culling due to inter-current illness or termination of the experiment) is reported as >x-y days.‡ Mean incubation periods for all three isolates are significantly lower in Tg152c hemizygotes than in Tg190 × Tg152c heterozygotes ( $P < 0.0001$ ; two-tailed unpaired t-test).**Table 3 | Transmission of human prions to transgenic mice homozygous for human PrP V127**

| Inoculum  |             |                               | Transmission data       |                               |                         |                               |
|-----------|-------------|-------------------------------|-------------------------|-------------------------------|-------------------------|-------------------------------|
| Aetiology | Source code | Human PrP <sup>Sc</sup> type* | Attack rate†            | Incubation period (days p.i.) | Attack rate†            | Incubation period (days p.i.) |
|           |             |                               | Tg183 V127M129/V127M129 |                               | Tg190 V127M129/V127M129 |                               |
| Kuru      | I516        | T3 VV                         | 0/19                    | >488–604                      | 0/22                    | >391–609                      |
| Kuru      | I520        | T3 VV                         | 0/19                    | >391–609                      | 0/19                    | >370–609                      |
| Kuru      | I10336      | T3 MV                         | 0/21                    | >405–607                      | 0/24                    | >463–617                      |
| Kuru      | I518        | T2 MM                         | 0/12                    | >439–600                      | 0/13                    | >432–603                      |
| vCJD      | I342        | T4 MM                         | 0/8                     | >553–605                      | 0/10                    | >506–604                      |
| vCJD      | I7042       | T4 MM                         | 0/11                    | >522–609                      | 0/14                    | >446–607                      |
| iCJD (GH) | I035        | T1 MM                         | 0/10                    | >434–622                      | 0/10                    | >450–602                      |
| sCJD      | I11058      | T1 MM                         | 0/8                     | >411–609                      | 0/9                     | >454–612                      |
| iCJD (DM) | I026        | T2 MM                         | 0/9                     | >381–602                      | 0/8                     | >516–602                      |
| sCJD      | I7040       | T2 MM                         | 0/8                     | >524–601                      | 0/6                     | >564–600                      |
| sCJD      | I280        | T2 VV                         | 0/9                     | >532–602                      | 0/9                     | >425–601                      |
| sCJD      | I278        | T2 VV                         | 0/7                     | >530–602                      | 0/7                     | >417–601                      |
| sCJD      | I284        | T2 MV                         | 0/6                     | >549–603                      | 0/6                     | >466–609                      |
| sCJD      | I1478       | T2 MV                         | 0/7                     | >517–602                      | 0/7                     | >418–617                      |
| sCJD      | I7394       | T3 VV                         | 0/8                     | >500–600                      | 0/8                     | >600–603                      |
| sCJD      | I764        | T3 MV                         | 0/10                    | >447–602                      | 0/8                     | >484–607                      |
| iCJD (GH) | I2651       | T3 VV                         | 0/7                     | >563–606                      | 0/7                     | >487–599                      |
| iCJD (GH) | I020        | T3 MV                         | 0/6                     | >546–603                      | 0/6                     | >572–600                      |

p.i., days post inoculation; vCJD, variant CJD; sCJD, sporadic CJD; iCJD, iatrogenic CJD; GH, growth hormone; DM, dura mater.

\* According to classification of ref. 28.

† Attack rate is defined as the total of clinically affected and subclinically infected mice as a proportion of the number of inoculated mice. Subclinical prion infection was assessed by immunoblotting and/or immunohistochemical examination of brain. As no clinical transmission of prion disease was observed, the interval between inoculation and death (from either senescence, culling due to inter-current illness or termination of the experiment) is reported as &gt;x-y days.

propagation. Here we demonstrate that HuPrP V127 is intrinsically resistant to prion conversion and indeed capable of inhibiting propagation of wild-type prions in a dose-dependent manner. Mice expressing only HuPrP V127 appear to be as resistant to prion disease as PrP null mice<sup>26</sup>, and understanding the structural basis of this effect may therefore provide critical insight into the molecular mechanism of mammalian prion propagation.

At its height, when first recognized by Western medicine in 1957, kuru was a devastating epidemic largely affecting women and children and killing up to 2% of the population annually in some villages. Some villages became largely devoid of young women of childbearing age. While collapse of the Fore population was prevented by cessation of endocannibalism in the late 1950s, which interrupted the route of transmission and led to a gradual decline in incidence<sup>27</sup>, our data suggest that if transmission had continued the epicentre of the affected region might have been repopulated with kuru-resistant individuals as a population genetic response to the epidemic. Understanding the structural basis of why HuPrP V127 is unable to propagate prions of multiple strain types may provide key insights into prion propagation and the development of rational therapeutics.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 28 January; accepted 7 May 2015.**

**Published online 10 June 2015.**

- Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24**, 519–550 (2001).
- Mead, S. *et al.* A novel protective prion protein variant that colocalizes with kuru exposure. *N. Engl. J. Med.* **361**, 2056–2065 (2009).
- Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).
- Collinge, J., Palmer, M. S. & Dryden, A. J. Genetic predisposition to iatrogenic Creutzfeldt-Jakob disease. *Lancet* **337**, 1441–1442 (1991).
- Palmer, M. S., Dryden, A. J., Hughes, J. T. & Collinge, J. Homozygous prion protein genotype predisposes to sporadic Creutzfeldt-Jakob disease. *Nature* **352**, 340–342 (1991).
- Collinge, J. Molecular neurology of prion disease. *J. Neurol. Neurosurg. Psychiatry* **76**, 906–919 (2005).
- Mead, S. *et al.* Balancing selection at the prion protein gene consistent with prehistoric kuru-like epidemics. *Science* **300**, 640–643 (2003).
- Antonyuk, S. V. *et al.* Crystal structure of human prion protein bound to a therapeutic antibody. *Proc. Natl Acad. Sci. USA* **106**, 2554–2558 (2009).
- Collinge, J. Variant Creutzfeldt-Jakob disease. *Lancet* **354**, 317–323 (1999).
- Collinge, J. & Clarke, A. A general model of prion strains and their pathogenicity. *Science* **318**, 930–936 (2007).
- Shibuya, S., Higuchi, J., Shin, R. W., Tateishi, J. & Kitamoto, T. Protective prion protein polymorphisms against sporadic Creutzfeldt-Jakob disease. *Lancet* **351**, 419 (1998).
- Asante, E. A. *et al.* BSE prions propagate as either variant CJD-like or sporadic CJD-like prion strains in transgenic mice expressing human prion protein. *EMBO J.* **21**, 6358–6366 (2002).
- Asante, E. A., Li, Y. G., Gowland, I., Jefferys, J. G. & Collinge, J. Pathogenic human prion protein rescues PrP null phenotype in transgenic mice. *Neurosci. Lett.* **360**, 33–36 (2004).
- Wadsworth, J. D. *et al.* Human prion protein with valine 129 prevents expression of variant CJD phenotype. *Science* **306**, 1793–1796 (2004).
- Wadsworth, J. D., Asante, E. A. & Collinge, J. Contribution of transgenic models to understanding human prion disease. *Neuropathol. Appl. Neurobiol.* **36**, 576–597 (2010).
- Collinge, J. *et al.* Unaltered susceptibility to BSE in transgenic mice expressing human prion protein. *Nature* **378**, 779–783 (1995).
- Hill, A. F. *et al.* The same prion strain causes vCJD and BSE. *Nature* **389**, 448–450 (1997).
- Asante, E. A. *et al.* Dissociation of pathological and molecular phenotype of variant Creutzfeldt-Jakob disease in transgenic human prion protein 129 heterozygous mice. *Proc. Natl Acad. Sci. USA* **103**, 10759–10764 (2006).
- Wadsworth, J. D. *et al.* Kuru prions and sporadic Creutzfeldt-Jakob disease prions have equivalent transmission properties in transgenic and wild-type mice. *Proc. Natl Acad. Sci. USA* **105**, 3885–3890 (2008).
- Collinge, J., Sidle, K. C., Meads, J., Ironside, J. & Hill, A. F. Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. *Nature* **383**, 685–690 (1996).
- Bruce, M. E. *et al.* Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent. *Nature* **389**, 498–501 (1997).
- Perrier, V. *et al.* Dominant-negative inhibition of prion replication in transgenic mice. *Proc. Natl Acad. Sci. USA* **99**, 13079–13084 (2002).
- Hizume, M. *et al.* Human prion protein (PrP) 219K is converted to PrP<sup>Sc</sup> but shows heterozygous inhibition in variant Creutzfeldt-Jakob disease infection. *J. Biol. Chem.* **284**, 3603–3609 (2009).
- Striebel, J. F., Race, B., Meade-White, K. D., LaCasse, R. & Chesebro, B. Strain specific resistance to murine scrapie associated with a naturally occurring human prion protein polymorphism at residue 171. *PLoS Pathog.* **7**, e1002275 (2011).
- DiSalvo, S., Derdowski, A., Pezza, J. A. & Serio, T. R. Dominant prion mutants induce curing through pathways that promote chaperone-mediated disaggregation. *Nature Struct. Mol. Biol.* **18**, 486–492 (2011).
- Büeler, H. *et al.* Mice devoid of PrP are resistant to scrapie. *Cell* **73**, 1339–1347 (1993).
- Alpers, M. P. The epidemiology of kuru: monitoring the epidemic from its peak to its end. *Phil. Trans. R. Soc. Lond. B* **363**, 3707–3713 (2008).
- Hill, A. F. *et al.* Molecular classification of sporadic Creutzfeldt-Jakob disease. *Brain* **126**, 1333–1346 (2003).

**Acknowledgements** This work was funded by the UK Medical Research Council. We thank all staff at our Biological Services Facility for animal care; C. O'Malley and S. Lyall for technical support with histology; and R. Newton and R. Young for preparation of figures. We thank patients and families for consent to use human tissues. We thank UK neurological and neuropathology colleagues for providing brain tissue from CJD patients. Some of this work was undertaken at University College London Hospitals/University College London, which received a proportion of funding from the Department of Health's National Institute for Health Research Biomedical Research Centres funding scheme. We thank the Fore communities and P. Siba and colleagues at the Papua New Guinea Institute of Medical Research for their long-term support. We thank the late C. Gajdusek, the late J. Gibbs, S. Landis, and their associates from the Laboratory of Central Nervous System Studies of the National Institutes of Health, Bethesda, USA, for archiving and sharing kuru samples.

**Author Contributions** E.A.A., J.D.F.W. and J.C. conceived and designed the study. M.S. and E.A.A. cloned the transgene constructs. R.H., A.T. and A.J. generated and developed transgenic lines. M.S., T.J., S.H. and E.A.A. analysed and characterized the transgenic mice. A.G. analysed prion-infected mouse brains by immunoblotting. A.R.-L., J.M.L. and S.B. performed histological analyses and interpreted the pathology data. S.M., J.W., J.D.F.W. and M.A. characterized CJD cases and inocula. E.A.A. supervised the study and collated data. E.A.A., J.D.F.W. and J.C. drafted the paper with contributions from all authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.C. ([j.collinge@prion.ucl.ac.uk](mailto:j.collinge@prion.ucl.ac.uk)).

## METHODS

No statistical methods were used to predetermine sample size.

**Ethics statement.** Storage and biochemical analyses of post-mortem human brain samples and transmission studies to mice were performed with written informed consent from patients with capacity to give consent. Where patients were unable to give informed consent, assent was obtained from their relatives in accordance with UK legislation and Codes of Practice. Samples were stored and used in accordance with the Human Tissue Authority Codes of Practice and in line with the requirements of the Human Tissue Authority licence held by UCL Institute of Neurology. This study was performed with approval from the Medical Research Advisory Committee of the Government of Papua New Guinea, the National Hospital for Neurology and Neurosurgery and the UCL Institute of Neurology Joint Research Ethics Committee (now National Research Ethics Service Committee, London – Queen Square); REC references: 03/N036, 03/N038 and 03/N133. Work with mice was performed under approval and licence granted by the UK Home Office (Animals (Scientific Procedures) Act 1986); Project Licence number 70/6454 which conformed to University College London institutional and ARRIVE guidelines (<http://www.nc3rs.org.uk/ARRIVE/>).

**Generation of transgenic mice.** The 759-bp human PrP ORF was amplified by PCR with pfu polymerase from genomic DNA prepared from the brain of a patient with the V127 polymorphism on the M129 allele, using forward primer 5'-GTCGACCAGTCATTATGGCGAACCTT-3' and reverse primer 5'-CTCGAGAAGACCTTCTCATCCCACT-3'. Restriction sites Sall (GTCGAC) and Xho I (CTCGAG) (underlined) were introduced in the forward and reverse primers respectively for cloning. The blunt-ended PCR fragment generated by pfu polymerase was subcloned into SmaI-digested pSP72 vector and sequenced to ensure that no spurious alterations had been introduced by the PCR and to confirm the presence of V127 and M129 polymorphisms matching the patient DNA template. The amplified human PrP ORF with the confirmed polymorphisms was then isolated by Sall and Xho I digestion. Subsequent subcloning into the Sall site of the cosmid vector SHaCosTt, packaging and preparation of high quality DNA of the Not I transgene insert was as previously reported<sup>12</sup>. Microinjection of the purified Not I DNA fragment was carried out according to standard protocol into single cell eggs of *Prnp* null mice which had been backcrossed onto an FVB/N genetic background. Genotyping was performed by PCR and PrP expression levels estimated by western blot analysis as previously reported<sup>12,29</sup>. Two homozygous lines were established for variant HuPrP V127M129 described as Tg(HuPrP V127M129/V127M129 *Prnp*<sup>0/0</sup>)-183 (V127M129 Tg183) and Tg(HuPrP V127M129/V127M129 *Prnp*<sup>0/0</sup>)-190 (V127M129 Tg190), with PrP<sup>C</sup> expression levels of 2 and 1 times, respectively, that of pooled 10% (w/v) normal human brain homogenate (Table 1). FVB-congenic versions of Tg35 mice homozygous for HuPrP G127M129 (refs 12–15) and Tg152 mice homozygous for HuPrP G127V129 (refs 15–19) were used as wild-type human PrP-expressing controls, designated Tg(HuPrP G127M129/G127M129 *Prnp*<sup>0/0</sup>)-35c (G127M129 Tg35c) and Tg(HuPrP G127V129/G127V129 *Prnp*<sup>0/0</sup>)-152c (G127V129 Tg152c), respectively. The *PRNP* codon 127–129 genotypes and relative expression levels of wild-type and variant PrP in the parental transgenic lines and in the F<sub>1</sub> crosses are shown in Table 1.

**Transmission studies.** Strict biosafety protocols were followed. Inocula were prepared, using disposable equipment for each inoculum, in a microbiological containment level 3 laboratory and inoculations performed within a class 1

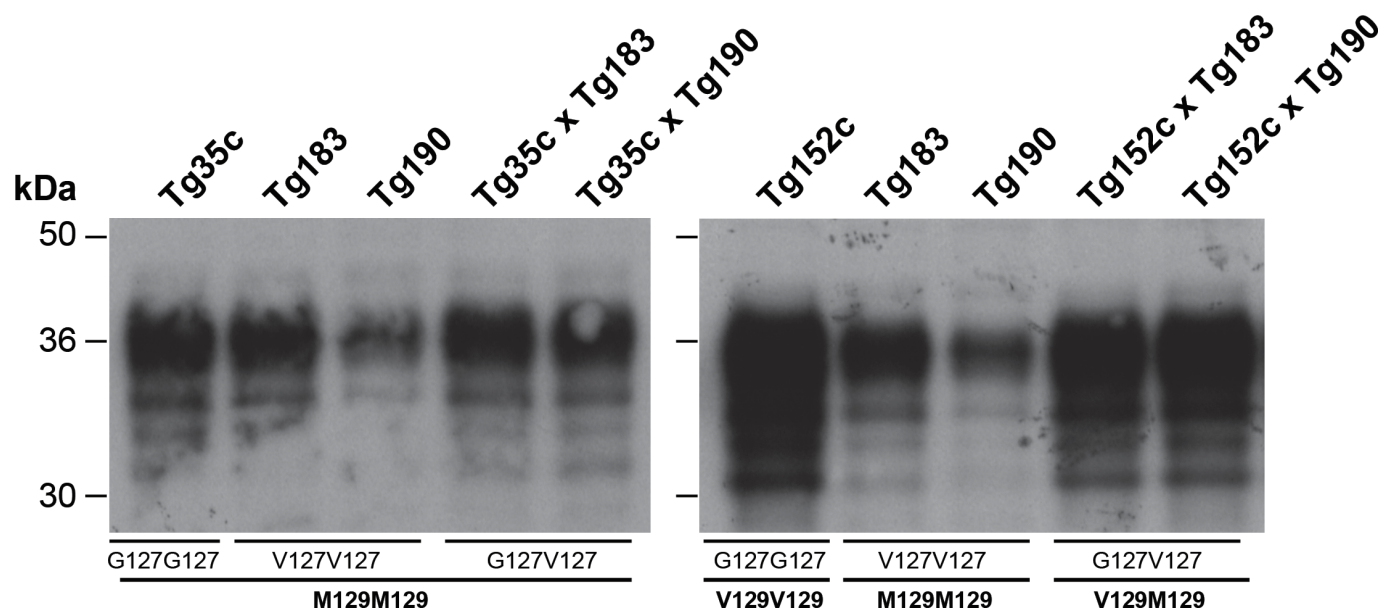
microbiological safety cabinet as described previously<sup>12,19</sup>. Ten mice per group of control G127M129 Tg35c and G127V129 Tg152c lines and 15 per group of newly generated V127M129 Tg183 and V127M129 Tg190 lines and their respective crosses with Tg35c and Tg152c were inoculated (see below) with human brain homogenates from neuropathologically confirmed patients comprising four kuru cases, eight sporadic CJD cases, four iatrogenic CJD cases and two cases of vCJD. The genotype of each mouse was confirmed by PCR of ear punch DNA before inclusion and all mice were uniquely identified by sub-cutaneous transponders. Disposable cages were used and all cage lids and water bottles were also uniquely identified by transponder and remained with each cage of mice throughout the incubation period. Care of the mice was according to institutional guidelines. Mice (female, aged 6–8 weeks) were randomly assigned to experimental groups and anaesthetized with a mixture of halothane and O<sub>2</sub>, and intracerebrally inoculated into the right parietal lobe with 30 µl of a 1% (w/v) brain homogenate prepared in Dulbecco's phosphate buffered saline lacking Ca<sup>2+</sup> or Mg<sup>2+</sup> ions (D-PBS). All mice were thereafter examined daily for early indicators of clinical prion disease including piloerection, sustained erect ears, intermittent generalized tremor, unsustained hunched posture, rigid tail, mild loss of coordination, and clasp hind legs when lifted by the tail. Definite diagnosis of clinical prion disease (triggering experimental end point) was reached if mice exhibited any two early indicator signs in addition to one confirmatory sign, or any two confirmatory signs. The confirmatory signs included ataxia, impairment of righting reflex, dragging of hind limbs, sustained hunched posture, or prominent abnormal breathing. Mice were killed (by CO<sub>2</sub> asphyxiation) if they exhibited any signs of distress or once a diagnosis of prion disease was established. At post-mortem, brains from inoculated mice were removed, divided sagittally with half frozen and half fixed in 10% buffered formal saline. Subsequent immunohistochemical or biochemical investigations were performed blind to sample provenance.

**Neuropathology and immunohistochemistry.** Fixed brain was immersed in 98% formic acid for 1 h and paraffin wax embedded. Serial sections of 4 µm thickness were pre-treated by boiling for 10 min in a low ionic strength buffer (2.1 mM Tris, 1.3 mM EDTA, 1.1 mM sodium citrate, pH 7.8) before exposure to 98% formic acid for 5 min. Abnormal PrP accumulation was examined using anti-PrP monoclonal antibody ICSM 35 (D-Gen Ltd, London) on a Ventana automated immunohistochemical staining machine (Ventana Medical Systems Inc., Tucson, Arizona) using proprietary secondary detection reagents (Ventana Medical Systems Inc.) before development with 3'3' diaminobenzidine tetrachloride as the chromogen<sup>30</sup>. Harris haematoxylin and eosin staining was done by conventional methods. Appropriate controls were used throughout.

**Immunoblotting.** Preparation of brain homogenates (10% (w/v) in D-PBS), proteinase K digestion (75 µg ml<sup>-1</sup> for 1 h at 37 °C) and subsequent immunoblotting was performed as described previously<sup>28,30</sup>. Blots were probed with anti-PrP monoclonal antibody ICSM 35 (D-Gen Ltd, London) in conjunction with an anti-mouse IgG-alkaline phosphatase conjugate and development in chemiluminescent substrate (CDP-Star; Tropix Inc.). Primary screening of brain homogenates was performed blind to sample identity.

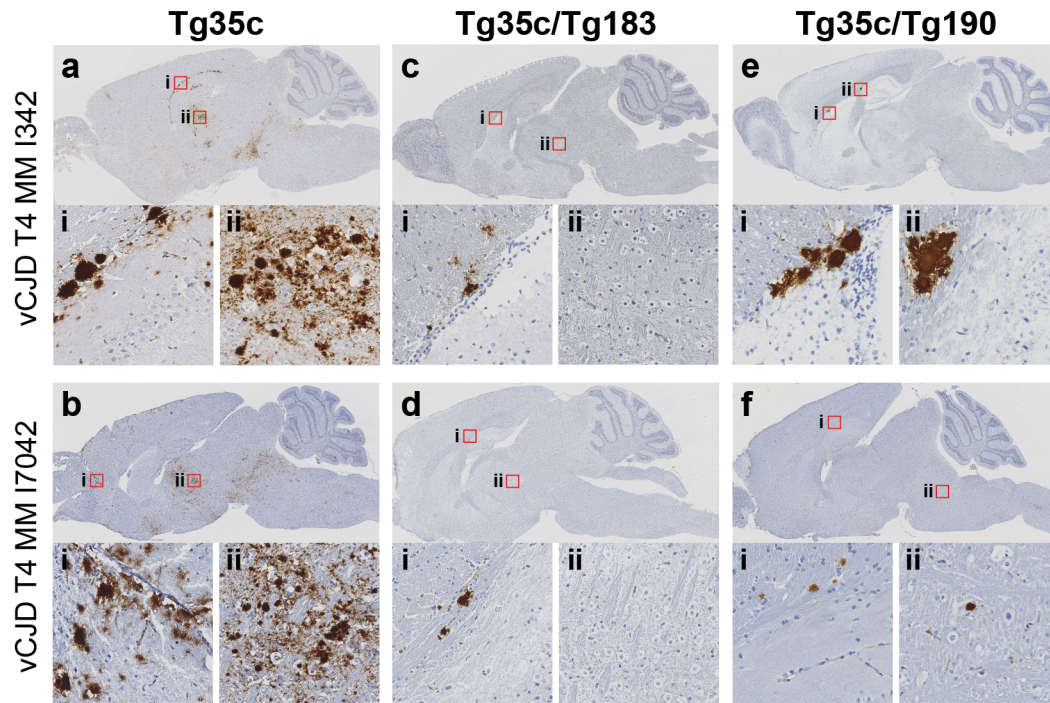
29. Asante, E. A. *et al.* Absence of spontaneous disease and comparative prion susceptibility of transgenic mice expressing mutant human prion proteins. *J. Gen. Virol.* **90**, 546–558 (2009).
30. Wadsworth, J. D. *et al.* Molecular diagnosis of human prion disease. *Methods Mol. Biol.* **459**, 197–227 (2008).





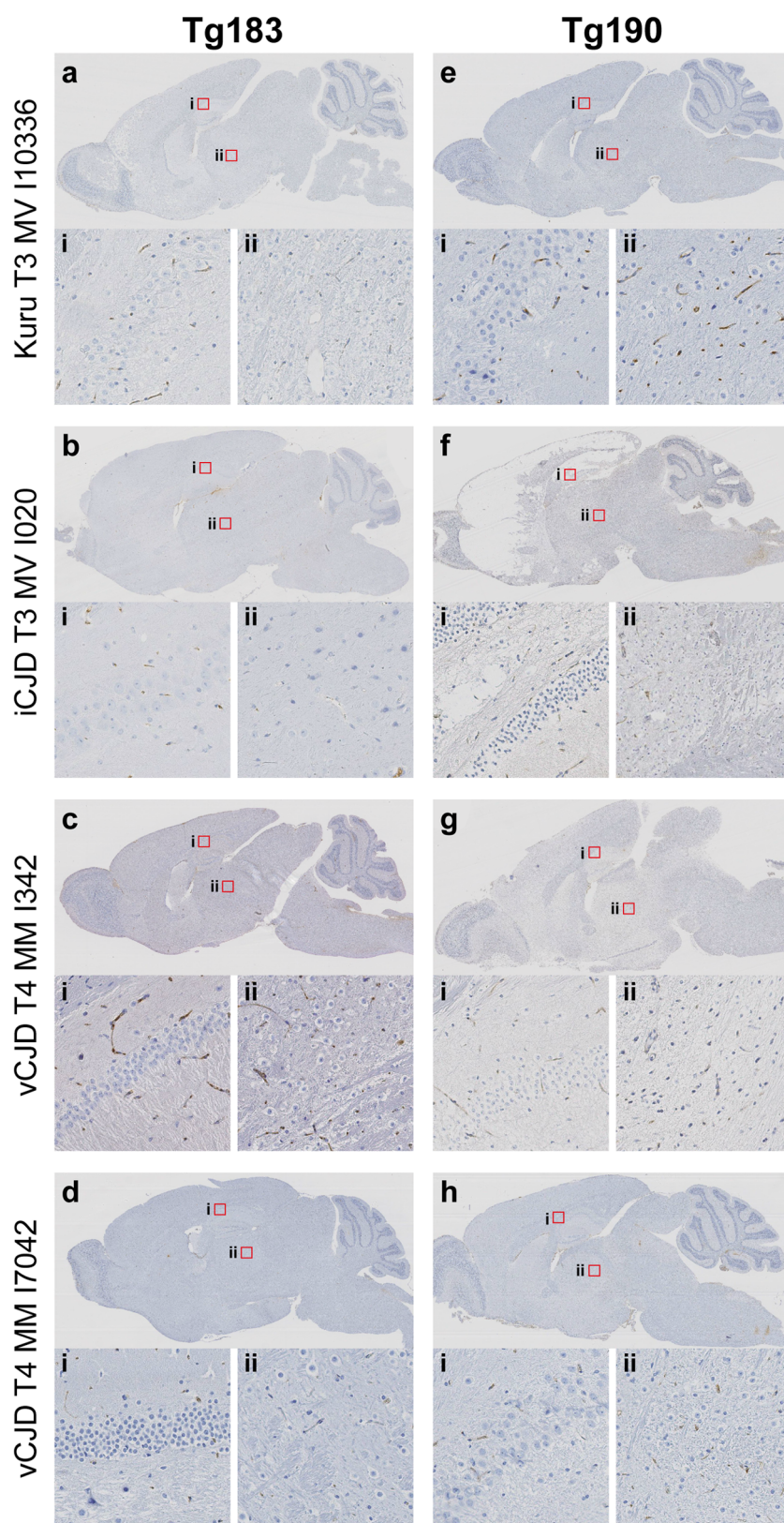
**Extended Data Figure 1 | Immunoblots showing relative PrP<sup>C</sup> expression levels in transgenic mice.** The provenance of each brain sample is designated above each lane and molecular markers are indicated on the left. The *PRNP* codon 127 (G, glycine, V, valine) or codon 129 (M, methionine, V, valine) genotypes of the transgenic mice are designated below. Transgenic mouse brains were analysed by enhanced chemiluminescence without proteinase-K digestion and equal amounts of total protein loaded in each well and probed

with anti-PrP monoclonal antibody ICSM 35. Wild-type human PrP expression levels of G127M129/G127M129 Tg35c and G127V129/G127V129 Tg152c mice are two- and sixfold higher, respectively, than seen in 10% (w/v) pooled human brain homogenate. Homozygous V127M129 Tg183 and Tg190 mice have twofold higher or equivalent PrP expression levels, respectively, compared to 10% (w/v) pooled human brain homogenate.



**Extended Data Figure 2 | Immunohistochemical analysis of transgenic mouse brain following challenge with vCJD prions.** All mice were intracerebrally challenged with the same vCJD prion isolates (I342 and I7042). Abnormal PrP deposition in fixed post-mortem brain from affected mice was detected using anti-PrP monoclonal antibody ICSM 35. **a, b**, Homozygous Tg35c mice (expressing G127M129/G127M129 wild-type PrP only) show intense and widespread PrP plaque deposits. Magnified areas show hippocampus (**a**, panel i), frontal cortex (**b**, panel i) and thalamus (**a**, panel ii, and **b**, panel ii). **c, d**, In contrast, heterozygous G127M129/V127M129 Tg35c/

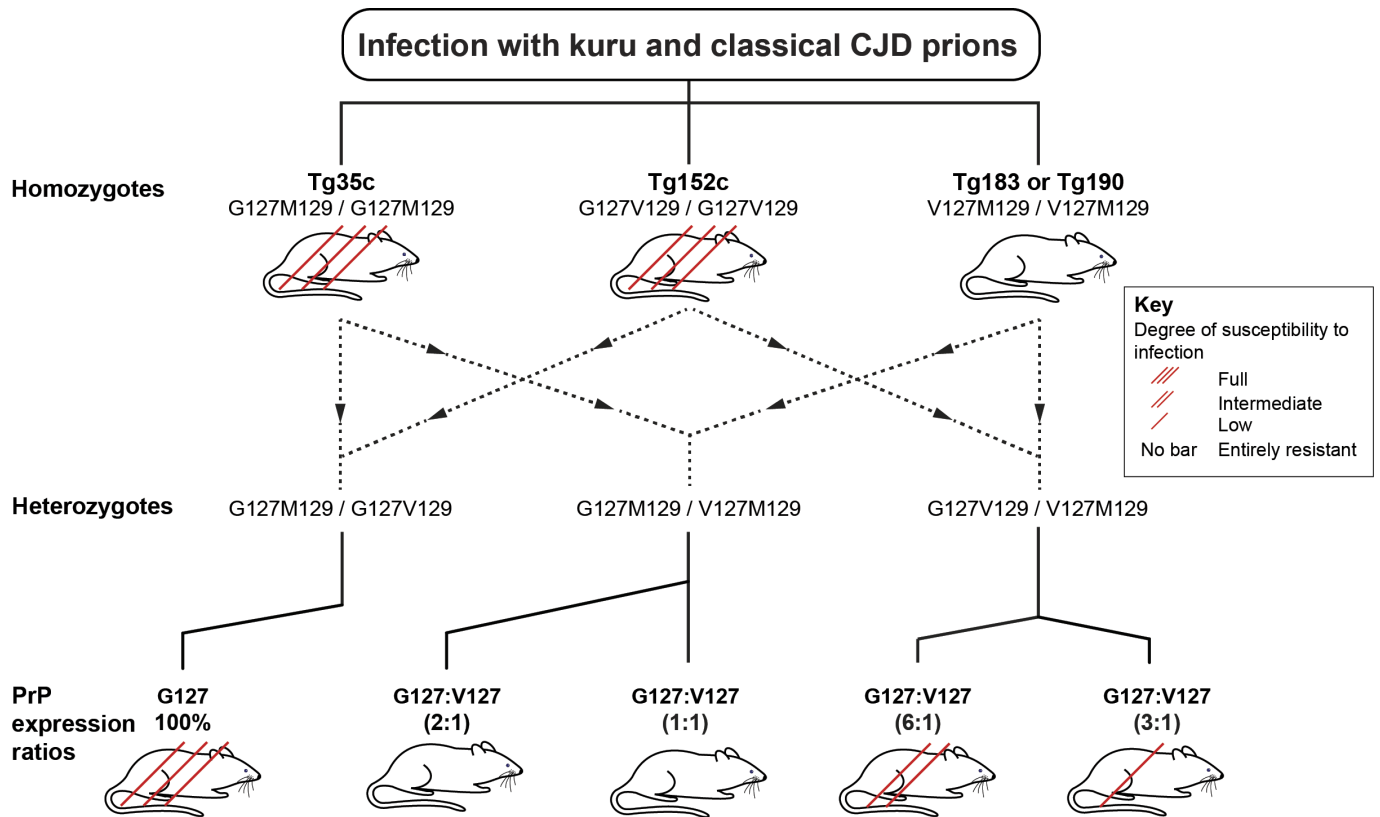
Tg183 mice expressing equivalent levels of G127M129 and V127M129 PrP show only weak PrP deposition in the corpus callosum (**c**, panel i, and **d**, panel i) with no abnormal PrP deposition detected in other brain areas, for example, in the thalamus (**c**, panel ii, and **d**, panel ii). **e, f**, Heterozygous G127M129/V127M129 Tg35c/Tg190 mice which express a lower level of V127M129 PrP relative to wild-type G127M129 PrP show greater levels of PrP deposition than seen in Tg35c/Tg183 mice following challenge with the same vCJD prion isolates. **e** (panels i and ii) and **f** (panel i), corpus callosum; **f** (panel ii), pons. Scale bar: upper panels, 2 mm; magnified panels, 100  $\mu$ m.



**Extended Data Figure 3 | Immunohistochemical analysis of homozygous *PRNP* V127M129/V127M129 transgenic mouse brain following challenge with human prions.** Mice were intracerebrally challenged with kuru, classical and variant CJD prions. Following prolonged (>600 days) post-inoculation periods, abnormal PrP deposition in fixed post-mortem brain was examined using anti-PrP monoclonal antibody ICSM 35. **a–d**, V127M129/V127M129 Tg183 mice with twofold overexpression of V127M129 PrP. **e–h**, V127M129/

V127M129 Tg190 mice expressing endogenous levels of V127M129 PrP. Red square boxes in the main panels (**a–h**) define the area of magnified images of hippocampus (panel **i**) and thalamus (panel **ii**) shown in the lower panels. Scale bars, 2 mm; magnified panels **i** and **ii**, 100  $\mu$ m. The lack of detection of abnormal PrP deposition in brain indicates that the mice are not subclinically infected with prions.

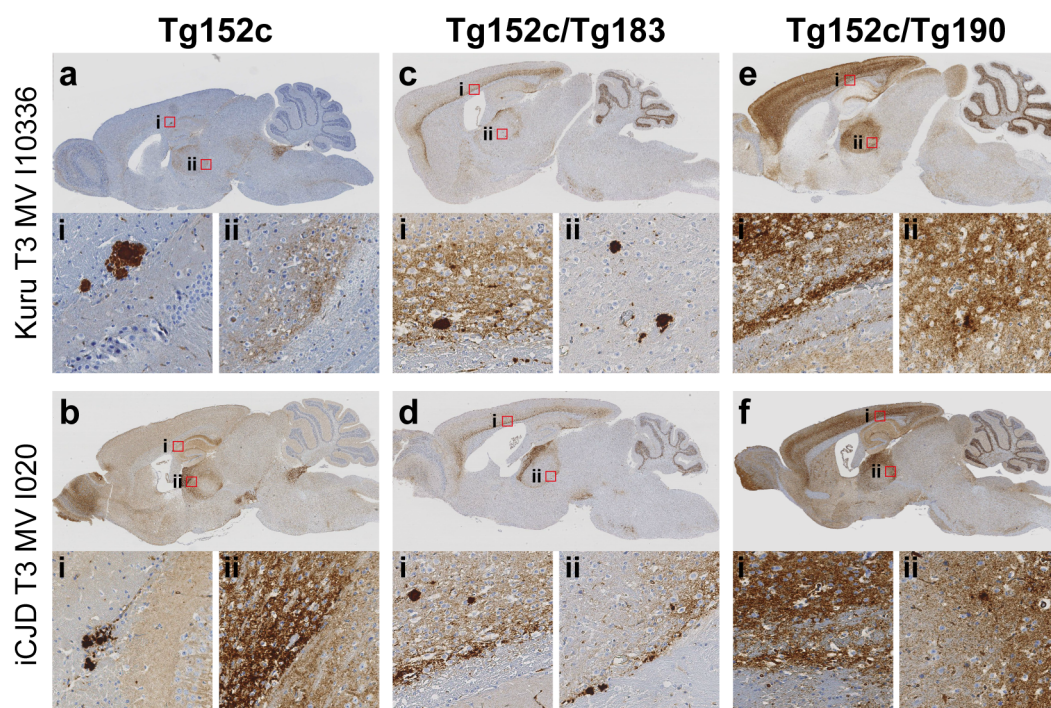




**Extended Data Figure 4 | Schematic diagram summarizing transmissions of kuru and classical CJD prions to transgenic mice expressing human PrP.** G127M129/G127M129 Tg35c mice or G127V129/G127V129 Tg152c mice are homozygous for wild-type human PrP alleles and are fully susceptible to kuru and classical CJD prions. V127M129/V127M129 Tg183 or V127M129/V127M129 Tg190 transgenic mice are homozygous for the variant V127M129 allele found only in humans from the kuru-exposed population of Papua New Guinea and are entirely resistant to infection with kuru and classical CJD prions. The levels of PrP expression in the brain of these homozygous transgenic mice relative to a pooled human brain homogenate are 1× (Tg190

mice), 2× (Tg35c and Tg183 mice) and 6× (Tg152c mice). Generation of F<sub>1</sub> mice through inter-breeding the various homozygous lines produces different combinations of the various human PrP alleles leading to differences in the relative expression levels of the various prion proteins in brain. The PrP expression ratios from the two PrP alleles in the crosses are shown in parentheses above the mouse sketches. Full, intermediate or low susceptibility of the mice to infection with kuru and classical CJD prions is indicated by three, two or one diagonal red bar, respectively, drawn across the mice. Mice with no red bar are entirely resistant to infection with kuru and classical CJD prions.





**Extended Data Figure 5 | Immunohistochemical analysis of homozygous *PRNP* G127V129/G127V129 and heterozygous G127V129/V127M129 transgenic mouse brain following challenge with human prions.** Mice were intracerebrally challenged with kuru and classical CJD prions, and abnormal PrP deposition in fixed post-mortem brain from affected mice was examined using anti-PrP monoclonal antibody ICSM 35. Red square boxes labelled i and

ii in panels a–f mark brain areas that are magnified and displayed below; i, cortex; ii, thalamus. **a, b**, Wild-type G127V129/G127V129 Tg152c mice. **c, d**, Heterozygous G127V129/V127M129 Tg152c/Tg183 mice. **e, f**, Heterozygous G127V129/V127M129 Tg152c/190 mice. Scale bars, 2 mm; magnified panels i and ii, 100  $\mu$ m. The detection of abnormal PrP deposition in brain indicates that the mice are infected with prions.

# The *Drosophila* TNF receptor Grindelwald couples loss of cell polarity and neoplastic growth

Ditte S. Andersen<sup>1,2,3,4\*</sup>, Julien Colombani<sup>1,2,3,4\*</sup>, Valentina Palmerini<sup>5</sup>, Krittalak Chakrabandhu<sup>1,2,3,6</sup>, Emilie Boone<sup>1,2,3,4</sup>, Michael Röthlisberger<sup>7</sup>, Janine Toggweiler<sup>7</sup>, Konrad Basler<sup>7</sup>, Marina Mapelli<sup>5</sup>, Anne-Odile Hueber<sup>1,2,3,6</sup> & Pierre Léopold<sup>1,2,3,4</sup>

**Disruption of epithelial polarity is a key event in the acquisition of neoplastic growth. JNK signalling is known to play an important part in driving the malignant progression of many epithelial tumours, although the link between loss of polarity and JNK signalling remains elusive. In a *Drosophila* genome-wide genetic screen designed to identify molecules implicated in neoplastic growth<sup>1</sup>, we identified *grindelwald* (*grnd*), a gene encoding a transmembrane protein with homology to members of the tumour necrosis factor receptor (TNFR) superfamily. Here we show that Grnd mediates the pro-apoptotic functions of Eiger (Egr), the unique *Drosophila* TNF, and that overexpression of an active form of Grnd lacking the extracellular domain is sufficient to activate JNK signalling *in vivo*. Grnd also promotes the invasiveness of *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> tumours through Egr-dependent Matrix metalloprotease-1 (Mmp1) expression. Grnd localizes to the subapical membrane domain with the cell polarity determinant Crumbs (Crb) and couples Crb-induced loss of polarity with JNK activation and neoplastic growth through physical interaction with Veli (also known as Lin-7). Therefore, Grnd represents the first example of a TNFR that integrates signals from both Egr and apical polarity determinants to induce JNK-dependent cell death or tumour growth.**

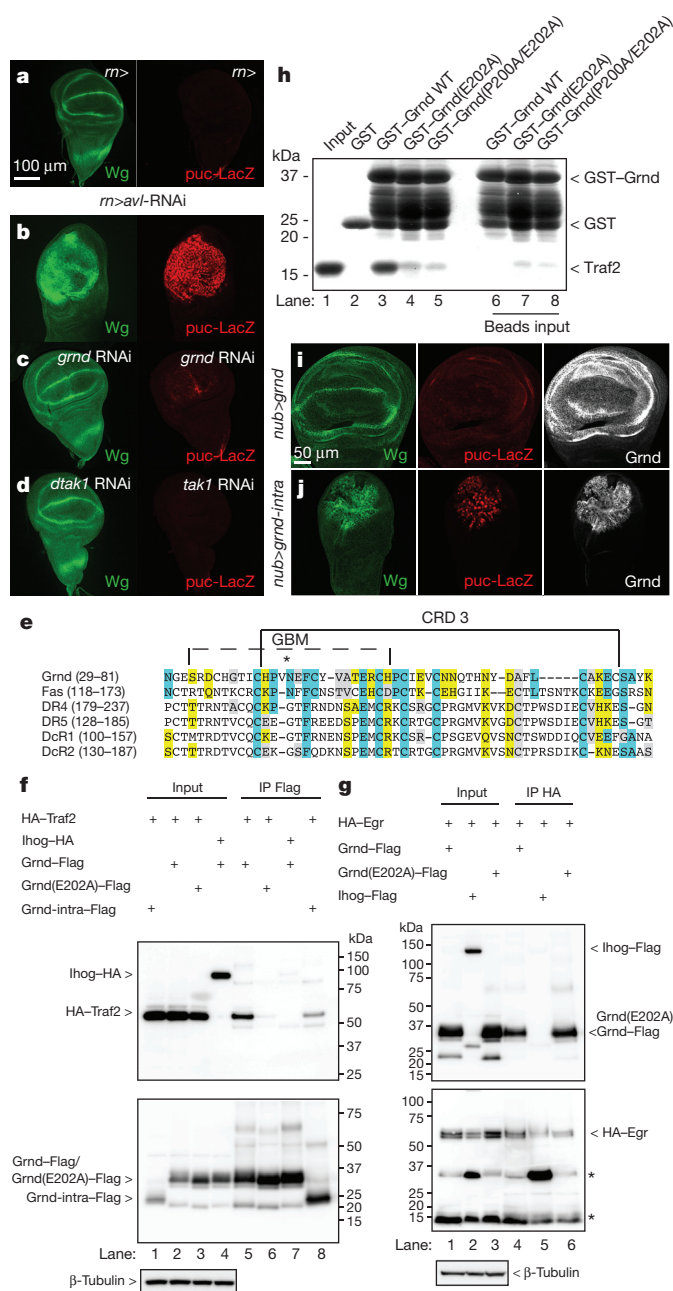
We recently carried out a genome-wide screen to identify molecules that are required for neoplastic growth<sup>1</sup>. The condition used for this screen was the disc-specific knockdown of *avalanche* (*rotund* (*rn*)>*avl*-RNAi; *avalanche* also known as *syntaxin 7*), a gene encoding a syntaxin that functions in the early step of endocytosis<sup>2</sup>. *rn*>*avl*-RNAi results in ectopic Wingless (Wg) expression, neoplastic disc overgrowth<sup>2</sup> (Fig. 1a, b), and a 2-day delay in larva-to-pupa transition<sup>1</sup>. We screened a collection of 10,100 transgenic RNA interference (RNAi) lines for their ability to rescue the pupariation delay and identified 121 candidate genes<sup>1</sup>. Interestingly, only eight candidate genes also rescued ectopic Wg expression and neoplastic overgrowth (Extended Data Fig. 1a). These included five lines targeting core components of the JNK pathway (Bendless, Tab2, Tak1, Hemipterous and Basket; Extended Data Fig. 1b). Using a *puckered* enhancer trap (*puc-lacZ*) as a readout for JNK activity, we confirmed that JNK signalling is highly upregulated in *rn*>*avl*-RNAi discs (Fig. 1a, b, right). One of the remaining lines targets *CG10176*, a gene encoding a transmembrane protein. Reducing expression of *CG10176* by using two different RNAi lines was as efficient as *tak1* silencing to restore normal Wg pattern and suppresses JNK signalling and neoplastic growth in the *rn*>*avl*-RNAi background (Fig. 1c, d and Extended Data Fig. 1c–e). Sequence analysis of *CG10176* identified a cysteine-rich domain (CRD) in the extracellular part with homology to vertebrate TNFRs (Fig. 1e and Extended Data Fig. 2) harbouring a glycosphingolipid-binding motif (GBM) characteristic of many TNFRs including Fas<sup>3</sup>

(Fig. 1e and Extended Data Fig. 2). We named *CG10176* *grindelwald* (*grnd*), after a village at the foot of Eiger, a Swiss mountain that lent its name to the unique *Drosophila* TNF, Egr. Immunostaining and subcellular fractionation of disc extracts confirmed that Grnd localizes to the membrane (Extended Data Fig. 1f–h). Moreover, co-immunoprecipitation experiments showed that both Grnd full-length and Grnd-intra, a form lacking its extracellular domain, directly associate with Traf2, the most upstream component of the JNK pathway (Fig. 1f–h). This interaction is disrupted by a single amino acid substitution within a conserved Traf6-binding motif<sup>4</sup> (human TRAF6 is the closest homologue to Traf2; Fig. 1f–h and Extended Data Fig. 3). Overexpression of Grnd-intra, but not full-length Grnd, is sufficient to induce JNK signalling, ectopic Wg expression and apoptosis (Fig. 1i–j and Extended Data Fig. 4a–c), and Grnd-intra-induced apoptosis is efficiently suppressed in a *hep*<sup>75</sup> (JNKK) mutant background (Extended Data Fig. 4d, e), confirming that Grnd acts upstream of the JNK signalling cascade.

The *Drosophila* TNF Egr activates JNK signalling and triggers cell death or proliferation, depending on the cellular context<sup>5</sup>. We therefore tested whether Grnd is required for the small-eye phenotype generated by Egr-induced apoptosis in the retinal epithelium<sup>6,7</sup> (*GMR*>*egr*; Fig. 2a, b). As previously shown, inhibition of JNK signalling by reducing *tak1* (ref. 6) or *traf2* (ref. 8) expression, or by overexpressing *puckered*<sup>7</sup>, blocks Egr-induced apoptosis and rescues the small-eye phenotype (Extended Data Fig. 5a–d). In contrast to a previous report<sup>9</sup>, RNAi silencing of *wengen* (*wgn*), a gene encoding a presumptive receptor for Egr, does not rescue the small-eye phenotype (Extended Data Fig. 5e). Furthermore, the small-eye phenotype is not modified in a *wgn*-null mutant background (Fig. 2c and Extended Data Fig. 5h, m, n), confirming that Wgn is not required for Egr-induced apoptosis in the eye. By contrast, reducing *grnd* levels partially rescues the Egr-induced small-eye phenotype, producing a ‘hanging-eye’ phenotype (Fig. 2d) that is not further rescued in a *wgn*-knockout (*wgn*<sup>KO</sup>) mutant background (Extended Data Fig. 5e–i). A similar phenotype was previously reported as a result of non-autonomous cell death induced by a diffusible form of Egr<sup>10</sup> (Extended Data Fig. 5k, l). This suggests that Grnd prevents Egr from diffusing outside of its expression domain. Co-immunoprecipitation experiments show that both full-length Grnd and Grnd-extra, a truncated form of Grnd lacking the cytoplasmic domain, associate with Egr through its TNF-homology domain (Fig. 2j, k). Although Grnd-extra can bind Egr, it cannot activate JNK signalling. Therefore, we reasoned that Grnd-extra expression might prevent both cell-autonomous and non-autonomous apoptosis by trapping Egr and preventing its diffusion and binding to endogenous Grnd. Indeed, *GMR*-Gal4-mediated expression of *grnd-extra* fully rescues the Egr small-eye phenotype (Fig. 2e and Extended Data Fig. 5j). To confirm that the removal of Grnd induces Egr-mediated non-autonomous cell death, we generated wing disc

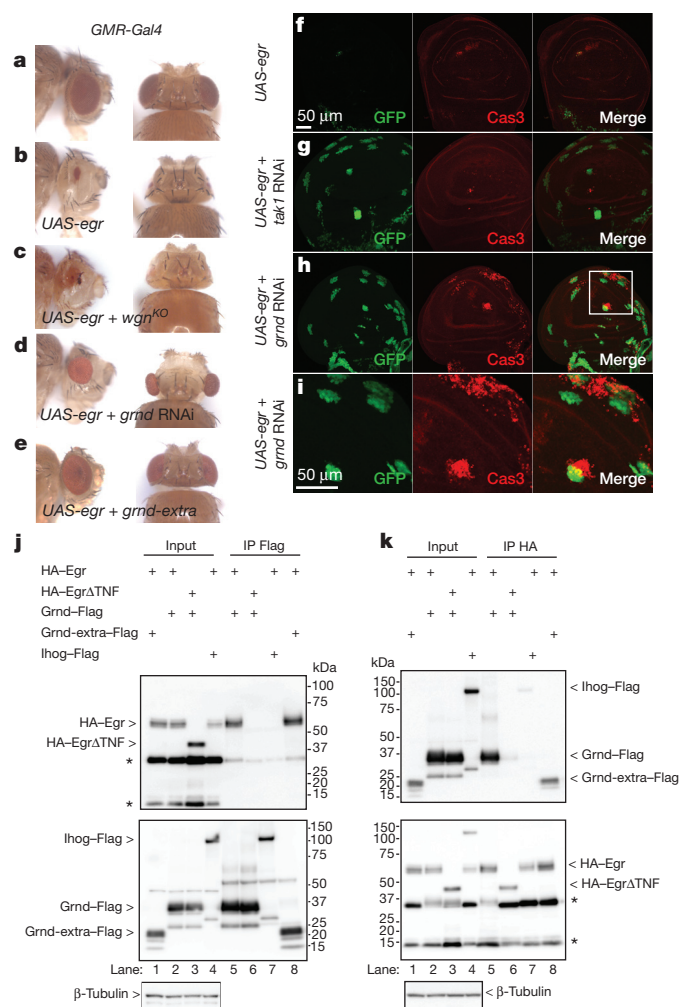
<sup>1</sup>University of Nice-Sophia Antipolis, Institute of Biology Valrose, Parc Valrose, 06108 Nice, France. <sup>2</sup>CNRS, Institute of Biology Valrose, Parc Valrose, 06108 Nice, France. <sup>3</sup>INSERM, Institute of Biology Valrose, Parc Valrose, 06108 Nice, France. <sup>4</sup>Genetics and Physiology of Growth laboratory, Institute of Biology Valrose, Parc Valrose, 06108 Nice, France. <sup>5</sup>Department of Experimental Oncology, European Institute of Oncology, Via Adamello 16, 20139 Milan, Italy. <sup>6</sup>Death receptors Signalling and Cancer Therapy laboratory, Institute of Biology Valrose, Parc Valrose, 06108 Nice, France. <sup>7</sup>Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

\*These authors contributed equally to this work.



**Figure 1 | Grnd is a novel member of the TNFR superfamily.** **a–d**, Wing discs stained for Wg (left) or LacZ (right). **e**, Sequence alignment of Grnd and human TNFRs. Blue, identity; yellow, strong similarity to Grnd; grey, weak similarity to Grnd. **f, g**, Co-immunoprecipitation experiments from S2R<sup>+</sup> cells co-expressing haemagglutinin (HA)–Traf2 or Ihog–HA (control), and either Grnd–Flag, Grnd-intra–Flag or Grnd(E202A)–Flag (**f**), or HA–Egr and either Grnd–Flag, Grnd(E202A)–Flag or Ihog–Flag (control) (**g**). Input loading is quantified using  $\beta$ -tubulin. **h**, Pull-down experiments of Traf2 with various truncated forms of glutathione S-transferase (GST)–Grnd. Coomassie blue staining is used to visualize the bound species. WT, wild type. **i–j**, Wing discs stained for Wg (left), LacZ (middle) and Grnd (right).

clones expressing *egr* alone, *egr* + *tak1* RNAi, or *egr* + *grmd* RNAi (Fig. 2f–i). As expected, reducing *tak1* levels in *egr*-expressing clones prevents their elimination by apoptosis (Fig. 2f, g). Similarly, reducing *grmd* levels prevents autonomous cell death, but also induces non-autonomous apoptosis (Fig. 2h, i). This suggests that Egr, like its mammalian counterpart TNF- $\alpha$ , can be processed into a diffusible form *in vivo* whose interaction with Grnd limits the potential to act

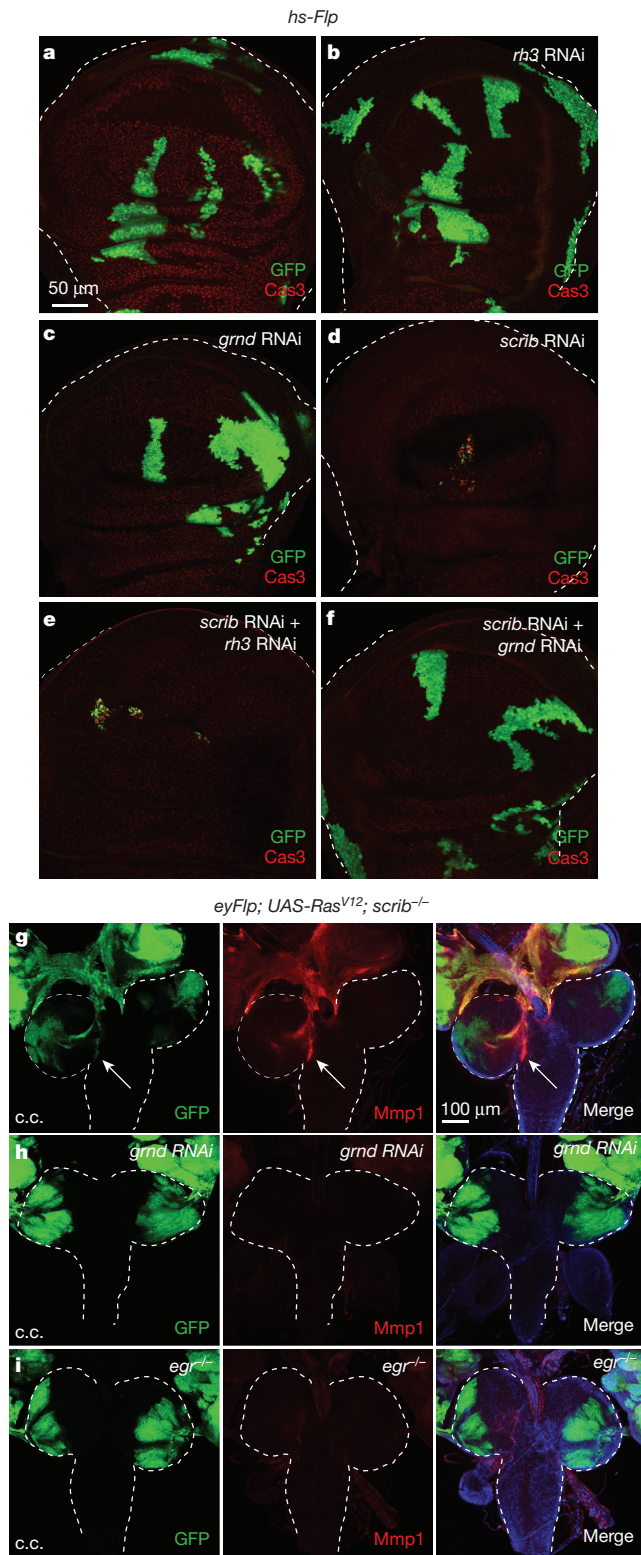


**Figure 2 | Grnd is required for Egr-induced apoptosis.** **a–e**, Light micrographs of *Drosophila* adult eyes of the indicated genotypes (all animals carry the *GMR-Gal4* driver). Original magnification,  $\times 30$ . **f–i**, Wing discs carrying green fluorescent protein (GFP)-labelled clones (green) expressing *egr* (**f**), *egr* + *tak1* RNAi (**g**) or *egr* + *grmd* RNAi (**h**, **i**) and stained for Caspase 3 (Cas3; red). A close-up shows non-autonomous cell death around the *egr* + *grmd* RNAi clones. **j, k**, Co-immunoprecipitation (IP) experiments from S2R<sup>+</sup> cells expressing HA–Egr or HA–Egr $\Delta$ TNF (Egr lacking its TNF-homology domain; amino acids 278–415) and either Grnd–Flag, Grnd-extra–Flag or Ihog–Flag (control). Input loading is quantified using  $\beta$ -tubulin.

at a distance. Flies carrying homozygous (*grmd*<sup>Minos/Minos</sup>) or transheterozygous (*grmd*<sup>Minos/Df</sup>) combinations of a transposon inserted in the *grmd* locus express no detectable levels of Grnd protein (Extended Data Fig. 6a, b) and are equally resistant to Egr-induced cell death (Extended Data Fig. 6e–j). In addition, *grmd*<sup>Minos/Minos</sup> mutant flies are viable and display no obvious phenotype (Extended Data Fig. 6c, d), suggesting that Grnd, like Egr, participates in a stress response to limit organismal damage. Collectively, our data demonstrate that Grnd is a new *Drosophila* TNF receptor that mediates most, if not all, Egr-induced apoptosis.

TNFs probably represent a danger signal produced in response to tissue damage to rid the organism of premalignant tissue or to facilitate wound healing. Disc clones mutant for the polarity gene *scribbled* (*scrib*) induce an Egr-dependent response resulting in the elimination of *scrib* mutant cells by JNK-mediated apoptosis<sup>11,12</sup>. To test the requirement for Grnd in this process, we compared *scrib*-RNAi and *scrib*-RNAi + *grmd*-RNAi clones obtained 72 h after heat shock induction. As expected, *scrib*-RNAi cells undergo apoptosis and detach from





**Figure 3 | Grnd is required for the Egr-induced invasiveness of *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> tumours.** a–f, Wing discs dissected 5 days after egg deposition (AED) stained for Cas3 (red) and carrying GFP-labelled clones (green) of the indicated genotypes induced 2 days AED. a, b, e, Wild-type clones (a) and clones expressing the control *Rh3* RNAi (b) to exclude a Gal4 titration effect (e). g–i, Eye–brain complexes dissected 7 days AED carrying *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> (g) or *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> + *grnd* RNAi (h) clones (labelled by GFP), or *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> GFP-labelled clones in an *egr* mutant background (i) stained for Mmp1 (g–i, middle, red) and 4',6-diamidino-2-phenylindole (DAPI; blue). *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> clones invade the ventral nerve cord (g, white arrow). c.c., cephalic complex.

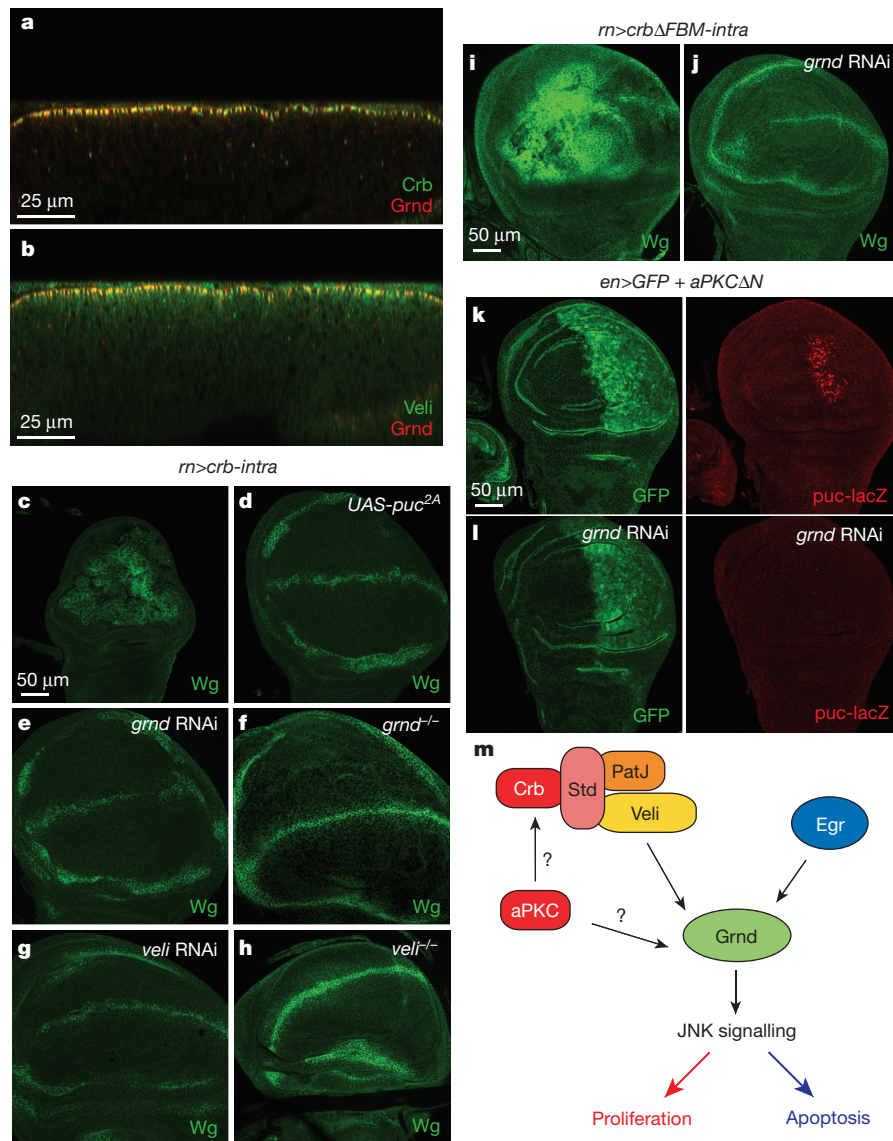
the epithelium (Fig. 3a–d). By contrast, *scrib*-RNAi clones with reduced *grnd* expression survive (Fig. 3e, f), indicating that Grnd is required for Egr-dependent elimination of *scrib*-RNAi cells. Similar results were obtained by generating *scrib*<sup>-/-</sup> mutant clones in the eye disc (Extended Data Fig. 7a–d).

In both mammals and flies, TNFs are double-edged swords that also have the capacity to promote tumorigenesis in specific cellular contexts<sup>5</sup>. Indeed, *scrib*<sup>-/-</sup> eye disc cells expressing an activated form of *Ras* (*Ras*<sup>V12</sup>) exhibit a dramatic tumour-like overgrowth and metastatic behaviour, a process that critically relies on Egr<sup>13</sup>. *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> metastatic cells show a strong accumulation of Grnd and Mmp1, and invade the ventral nerve cord<sup>11,14,15</sup> (Fig. 3g and Extended Data Figs 7e, 8a). Primary tumour cells reach peripheral tissues such as the fat body and the gut, where they form micro-metastases expressing high levels of Grnd (Extended Data Fig. 7f, g). Reducing *grnd* levels in *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> clones is sufficient to restore normal levels of Mmp1 and abolish invasiveness in a way similar to that observed in an *egr*<sup>-/-</sup> background<sup>13</sup> (Fig. 3h, i and Extended Data Fig. 8a–c). Therefore, Grnd is required for the Egr-induced metastatic behaviour of *Ras*<sup>V12</sup>/*scrib*<sup>-/-</sup> tumorous cells. Similarly, reducing *grnd*, but not *wgn* levels, strongly suppresses Mmp1 expression in *Ras*<sup>V12</sup>/*dlg*-RNAi cells and limits tumour invasion (Extended Data Fig. 8d–g), indicating that Wgn does not have a major role in the progression of these tumours.

Perturbation of cell polarity is an early hallmark of tumour progression in epithelial cells. In contrast to small patches of polarity-deficient cells, for example, *scrib*<sup>-/-</sup> clones, organ compartments or animals fully composed of polarity-deficient cells become refractory to Egr-induced cell death and develop epithelial tumours. The formation of these tumours requires JNK/MAPK signalling, but not Egr<sup>13,16</sup>, suggesting Egr-independent coupling between loss of polarity and JNK/MAPK-dependent tumour growth. In line with these observations, we noticed that, in contrast to Grnd, Egr is not required to drive neoplastic growth in *rn>avl*-RNAi conditions (Extended Data Fig. 9a, b). This suggests that, in addition to its role in promoting Egr-dependent functions, Grnd couples loss of polarity with JNK-dependent growth independently of Egr. Disc immunostainings revealed that Grnd co-localizes with the apical determinant Crb in the marginal zone, apical to the adherens junction protein E-cadherin (E-cad) and the atypical protein kinase C (aPKC; Fig. 4a and Extended Data Fig. 9c, d). In *avl*-RNAi discs, Grnd and Crb accumulate in a wider apical domain<sup>2</sup> (Extended Data Fig. 9e, g). Apical accumulation of Crb is proposed to be partly responsible for the neoplastic growth induced by *avl* knockdown, since overexpression of Crb or a membrane-bound cytoplasmic tail of Crb (Crb-intra) mimics the *avl*-RNAi phenotype<sup>2</sup> (Extended Data Fig. 9f). We therefore examined whether Grnd might couple the activity of the Crb complex with JNK-mediated neoplastic growth. Indeed, reducing *grnd* levels, but not *wgn*, in *rn>crb-intra* discs suppresses neoplastic growth as efficiently as inhibiting the activity of the JNK pathway (Fig. 4c–f and Extended Data Fig. 9h, i). Notably, Yki activation is not rescued in these conditions (Extended Data Fig. 9j–l), illustrating the ability of Crb-intra to promote growth independently of Grnd by inhibiting Hippo signalling through its FERM-binding motif (FBM)<sup>17,18</sup>. Indeed, neoplastic growth and polarity defects induced by a form of Crb-intra lacking its FBM (CrbΔFBM-intra) are both rescued by Grnd silencing (Fig. 4i, j). As expected, the size of *rn>crbΔFBM-intra;grnd*-RNAi discs is reduced compared to the size of *rn>crb-intra;grnd*-RNAi discs (compare Fig. 4e, f with Fig. 4j).

Crb, Stardust (Sdt; PALS1 in humans), and Pals1-associated tight junction protein (Patj) make up the core Crb complex<sup>19</sup>, which recruits the adaptor protein Veli (MALS1–3 in humans)<sup>20–22</sup>. In agreement with previous yeast two-hybrid data<sup>23</sup>, we find that Grnd binds directly and specifically to the PDZ domain of Veli through a membrane-proximal stretch of 28 amino acids in its intracellular domain (Extended Data Fig. 10a–e). Grnd localization is unaffected in





**Figure 4** | *grnd* is required downstream of the apical Crb polarity complex for JNK-dependent neoplastic growth. **a, b**, Transverse sections of wild-type discs stained for Grnd (red), Crb (**a**, green) or Veli (**b**, green). **c–h**, Discs expressing *crb-intra* alone (**c**), or combined with the indicated genotypes (**d–h**), dissected 6 days AED, stained for Wg (green). **i, j**, Discs expressing

*crb-intra* lacking its FBM (*crbΔFBM-intra*) alone (**i**) or with *grnd* RNAi (**j**) stained for Wg (in green). **k, l**, Discs from *puc-LacZ* animals expressing activated aPKC (aPKCΔN)<sup>26</sup> in the posterior domain alone (**k**) or together with *grnd* RNAi (**l**). **m**, A model for Grnd function.

*crb*<sup>-/-</sup> and *veli* RNAi mutant clones (Extended Data Fig. 10f–h). However, reducing *veli* expression rescues the patterning defects and disc morphology of *rn>crb-intra* mutant cells (Fig. 4g, h), suggesting that Grnd couples Crb activity with JNK signalling through its interaction with Veli. Interestingly, aPKC-dependent activation of JNK signalling<sup>16</sup> also depends on Grnd (Fig. 4k, l). aPKC is capable of directly binding and phosphorylating Crb, which is important for Crb function<sup>24</sup>. This suggests that aPKC, either directly or through Crb phosphorylation, activates Grnd-dependent JNK signalling in response to perturbation of apico-basal polarity.

Our data are consistent with a model whereby Grnd integrates signals from Egr, the unique fly TNF, and apical polarity determinants to induce JNK-dependent neoplastic growth or apoptosis in a context-dependent manner (Fig. 4m). Recent work reveals a correlation between mammalian Crb3 expression and tumorigenic potential in mouse kidney epithelial cells<sup>25</sup>. The conserved nature of the Grnd receptor suggests that specific TNFRs might carry out similar functions in vertebrates, in which the link between apical cell polarity and tumour progression remains elusive.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 September 2013; accepted 9 February 2015.

Published online 15 April 2015.

- Colombani, J., Andersen, D. S. & Leopold, P. Secreted peptide Dilp8 coordinates *Drosophila* tissue growth with developmental timing. *Science* **336**, 582–585 (2012).
- Lu, H. & Bilder, D. Endocytic control of epithelial polarity and proliferation in *Drosophila*. *Nature Cell Biol.* **7**, 1232–1239 (2005).
- Chakrabandhu, K. *et al.* The extracellular glycosphingolipid-binding motif of Fas defines its internalization route, mode and outcome of signals upon activation by ligand. *Cell Death Differ.* **15**, 1824–1837 (2008).
- Ye, H. *et al.* Distinct molecular mechanism for initiating TRAF6 signalling. *Nature* **418**, 443–447 (2002).
- Vidal, M. The dark side of fly TNF: an ancient developmental proof reading mechanism turned into tumor promoter. *Cell Cycle* **9**, 3851–3856 (2010).
- Igaki, T. *et al.* Eiger, a TNF superfamily ligand that triggers the *Drosophila* JNK pathway. *EMBO J.* **21**, 3009–3018 (2002).

7. Moreno, E., Yan, M. & Basler, K. Evolution of TNF signaling mechanisms: JNK-dependent apoptosis triggered by Eiger, the *Drosophila* homolog of the TNF superfamily. *Curr. Biol.* **12**, 1263–1268 (2002).
8. Xue, L. *et al.* Tumor suppressor CYLD regulates JNK-induced cell death in *Drosophila*. *Dev. Cell* **13**, 446–454 (2007).
9. Kanda, H., Igaki, T., Kanuka, H., Yagi, T. & Miura, M. Wengen, a member of the *Drosophila* tumor necrosis factor receptor superfamily, is required for Eiger signaling. *J. Biol. Chem.* **277**, 28372–28375 (2002).
10. Narasimamurthy, R. *et al.* Structure–function analysis of Eiger, the *Drosophila* TNF homolog. *Cell Res.* **19**, 392–394 (2009).
11. Brumby, A. M. & Richardson, H. E. *scribble* mutants cooperate with oncogenic Ras or Notch to cause neoplastic overgrowth in *Drosophila*. *EMBO J.* **22**, 5769–5779 (2003).
12. Igaki, T., Pastor-Pareja, J. C., Aonuma, H., Miura, M. & Xu, T. Intrinsic tumor suppression and epithelial maintenance by endocytic activation of Eiger/TNF signaling in *Drosophila*. *Dev. Cell* **16**, 458–465 (2009).
13. Cordero, J. B. *et al.* Oncogenic Ras diverts a host TNF tumor suppressor activity into tumor promoter. *Dev. Cell* **18**, 999–1011 (2010).
14. Pagliarini, R. A. & Xu, T. A genetic screen in *Drosophila* for metastatic behavior. *Science* **302**, 1227–1231 (2003).
15. Uhlirova, M. & Bohmann, D. JNK- and Fos-regulated Mmp1 expression cooperates with Ras to induce invasive tumors in *Drosophila*. *EMBO J.* **25**, 5294–5304 (2006).
16. Sun, G. & Irvine, K. D. Regulation of Hippo signaling by Jun kinase signaling during compensatory cell proliferation and regeneration, and in neoplastic tumors. *Dev. Biol.* **350**, 139–151 (2011).
17. Ling, C. *et al.* The apical transmembrane protein Crumbs functions as a tumor suppressor that regulates Hippo signaling by binding to Expanded. *Proc. Natl Acad. Sci. USA* **107**, 10532–10537 (2010).
18. Robinson, B. S., Huang, J., Hong, Y. & Moberg, K. H. Crumbs regulates Salvador/Warts/Hippo signaling in *Drosophila* via the FERM-domain protein Expanded. *Curr. Biol.* **20**, 582–590 (2010).
19. Bulgakova, N. A. & Knust, E. The Crumbs complex: from epithelial-cell polarity to retinal degeneration. *J. Cell Sci.* **122**, 2587–2596 (2009).
20. Bachmann, A., Grawe, F., Johnson, K. & Knust, E. *Drosophila* Lin-7 is a component of the Crumbs complex in epithelia and photoreceptor cells and prevents light-induced retinal degeneration. *Eur. J. Cell Biol.* **87**, 123–136 (2008).
21. Bachmann, A. *et al.* Cell type-specific recruitment of *Drosophila* Lin-7 to distinct MAGUK-based protein complexes defines novel roles for Sdt and Dlg-S97. *J. Cell Sci.* **117**, 1899–1909 (2004).
22. Olsen, O. *et al.* Renal defects associated with improper polarization of the CRB and DLG polarity complexes in MALS-3 knockout mice. *J. Cell Biol.* **179**, 151–164 (2007).
23. Giot, L. *et al.* A protein interaction map of *Drosophila melanogaster*. *Science* **302**, 1727–1736 (2003).
24. Sotillos, S., Diaz-Meco, M. T., Caminero, E., Moscat, J. & Campuzano, S. DaPKC-dependent phosphorylation of Crumbs is required for epithelial cell polarity in *Drosophila*. *J. Cell Biol.* **166**, 549–557 (2004).
25. Karp, C. M. *et al.* Role of the polarity determinant crumbs in suppressing mammalian epithelial tumor progression. *Cancer Res.* **68**, 4105–4115 (2008).
26. Betschinger, J., Mechtler, K. & Knoblich, J. A. The Par complex directs asymmetric cell division by phosphorylating the cytoskeletal protein Lgl. *Nature* **422**, 326–330 (2003).

**Acknowledgements** We thank G. Jarretou for technical assistance, the Vienna *Drosophila* RNAi Centers, the *Drosophila* Genetics Resource Center, the Bloomington Stock Center, N. Caridi and S. Pasqualato for technical assistance and P.L. laboratory members for comments on the manuscript. This work was supported by the CNRS, INSERM, Agence Nationale de la Recherche, Fondation pour la Recherche Médicale, Association pour la Recherche contre le Cancer (grant no. PJA20131200042 to J.C.), European Research Council (Advanced grant no. 268813 to P.L.), Marie Curie Life Long Training (grant no. 252373 to D.S.A.), the Labex Signallife program (grant ANR-11-LABX-0028-01 to P.L.), the Italian Association for Cancer Research (AIRC IG-12877) and the Italian Ministry of Health (GR-2008-1134103) to M.M.

**Author Contributions** D.S.A., J.C., K.C., A.-O.H., K.B., M.M. and P.L. designed the research; D.S.A., J.C., K.C., V.P., E.B., M.R. and J.T. performed experiments; D.S.A., J.C., K.C., E.B., A.-O.H., K.B., V.P., M.M. and P.L. analysed the data; D.S.A., J.C. and P.L. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.C. ([julien.colombani@unice.fr](mailto:julien.colombani@unice.fr)) or P.L. ([leopold@unice.fr](mailto:leopold@unice.fr)).

## METHODS

**Fly strains and food.** Animals were reared at 25 °C on fly food containing, per litre: 17 g inactivated yeast powder, 83 g corn flour, 10 g agar, 60 g white sugar and 4.6 g Nipagin M (in ethanol). The *rn-Gal4*, *UAS-egr*, *tubGal80*<sup>18</sup> animals were raised at 18 °C, shifted to 30 °C for 40 h during early third instar larval development (7 days AEL), and subjected to dissection immediately after (Extended Data Fig. 6).

**Cell culture.** *Drosophila* S2 and S2R<sup>+</sup> cells were grown in complete Schneider's medium (CSM; Schneiders medium (Invitrogen) supplemented with 10% heat-inactivated FBS (BioWhittaker), 50 U ml<sup>-1</sup> penicillin and 50 µg ml<sup>-1</sup> streptomycin (Invitrogen)) at 25 °C. Transfections were done using Effectene reagent (Qiagen).

**Plasmids.** *grnd*, *grnd-extra* and *grnd-intra* coding sequences were PCR amplified from BDGP EST complementary DNA clones RE28509, and cloned into the pENTR/D-TOPO vector using the following gene-specific primers: sense primers CACCATGTCGGTCAGGAAGTTGAG (*grnd* and *grnd-extra*) and CACCA TGCTGATCCTCTGCTGCTGACTATCC (*grnd-intra*); and antisense primers GAAAGCGACGGGAATCGTCGC (*grnd* and *grnd-intra*) and CCATCGAAGG AACTGAAACGCG (*grnd-extra*). The mutations giving rise to the single E202A and double P200A/E202A amino acid changes in Grnd and Grnd<sup>125–241</sup> were introduced using the QuickChange site-directed mutagenesis kit (Stratagene), *grnd-pENTR* or *grnd*<sup>125–241</sup>-pGEX-6PI templates, and the following primers: sense, CGGTGAGCACTCGCTATCCGGCGGCAACAGTACCACGCCAC; and antisense, GTGGGCGTGGTACTGTTTGGCGCGGATAGCGAGTGCTCAC CG for *grnd*<sup>E202A</sup>; sense, CGGTGAGCACTCGCTATGCGGCGGCAACAGT ACCACGCCAC; and antisense, GTGGGCGTGGTACTGTTTGGCGCGGCA TAGCGAGTGCTCACCG for *grnd*<sup>P200A/E202A</sup>. To generate *grnd-Flag*, *grnd*<sup>E202A</sup>-*Flag*, *grnd-extra-Flag* and *grnd-intra-Flag*, *grnd*, *grnd*<sup>E202A</sup>, *grnd-extra* and *grnd-intra* coding sequences were cloned into the Gateway Destination vector pActin-3×Flag (*Drosophila* Gateway Vector Collection) for expression in S2 and S2R<sup>+</sup> cells. In addition, *grnd*, *grnd-intra* and *grnd-extra* coding sequences were cloned into pUASg.attB for generation of UAS-*grnd*, UAS-*grnd-intra* and UAS-*grnd-extra* transgenic lines. The *HA-traf2* plasmid was from K.B.<sup>27</sup>. The *HA-egr* and *HA-egrΔTNF* plasmids were a gift from M. Miura. pActin-iHog-*HA* and pActin-iHog-*Flag* were gifts from L. Ruel.

**Transgenic flies.** To generate the UAS-*grnd* UAS-*grnd-intra* and UAS-*grnd-extra* lines, the pUASattB-*grnd*, pUASattB-*grnd-intra* and pUASattB-*grnd-extra* constructs were introduced into the germ line by injections in the presence of the PhiC31 integrase and inserted in the 86F8-landing site on the 3R chromosome (Bloomington *Drosophila* Stock Center, BL24749, BestGene). The following RNAi lines were from the GD or KK collections of the Vienna *Drosophila* RNAi Center (VDRC): *grnd* RNAi (GD43454, sense primer CGCGAATTCCTCTTGCTG CTGACTATCCTTCTCGT, and antisense primer GCGTCTAGAGTTGTGG GCGTGTACTGTTTTC, and KK104538, sense primer TGGCGCCCTAG ATGAACGGTTATGACGGAGATCG, and antisense primer CGCATGTAGC CTGCCGTTTGTGGTGATTGCGAATG), *tak1* RNAi (KK101357, sense primer TGGCGCCCTAGATGGCTTGTCTTGAAGTCTGTC, and antisense primer CGACGCCCGCTGATAGCATGAGCAGGAGATTGTCA), *bendless* RNAi (KK109638, sense primer TGGCGCCCTAGATGCGCATCATCAAGGAGA CTCA, and antisense primer CGCATGTAGCCTGCCTTCTGAAAAACGCC GATAC), *tab2* RNAi (KK100326, sense primer TGGCGCCCTAGATGGCA CTTTCAGCGGAAGAAAC, and antisense primer TAGCCTCCCTAGCG CAGCTCTGCACTGGGATAGGA), *traf2* RNAi (GD16125, sense primer CGC GAATTCAGACCCCACTCCGCCAAGA, and antisense primer GCGTCTAG AACCGAACAGCCAGCGAGGA), *hep* RNAi (KK109277, sense primer TGG CGCCCTAGATGGGCTTCTTGGACAGCTTGAG, and antisense primer CG CATGTAGCCTGCCAGGAAACAAACGCATCCTG), *bsk* RNAi (KK104569, sense primer TGGCGCCCTAGATGAGCATCCACTTCTCAGCAT, and antisense primer TAGGTCTAGCCCGCGGTACACCATCACCCTCGTT), *scrib* RNAi (GD 45556, sense primer CGCGAATTCGGAAGAGGACAACG CACAGCA, and antisense primer GCGTCTAGACTGGGAGCGGGCAAAGT GTAAG), *avl* (GD5413, sense primer CGCGAATTCACAGTAGCGCCAG CCAGCA, and antisense primer CGCTCTAGACCGTGCCTGCGAGACA AAG), *dilp8* (KK102604, sense primer TGGCGCCCTAGATGATGATGAC GACCTTCTGTGCG, and antisense primer CGCATGTAGCCTGCCGTATCG GAGTCTGTTGCCCT), and *rh3* RNAi (KK100853, sense primer TGGCGCCCC TAGATGATCTGGGATGACTTATGGCG, and antisense primer TAGCCTCC CTAGCGCTCGATCGCACTCTTGTG). The *veli*-RNAi lines were from U. Thomas. The *wengen*-RNAi line (*wengen-IR*) was provided by M. Miura. The *GMR-Gal4* (on chromosome II), *UAS-egr* (on chromosome III), *GMR-Gal4*, *UAS-ecto-egr60* (on chromosome II), *GMR-Gal4*, *UAS-egr* (on chromosome II), *UAS-CrbΔFBM-intra*, *DIAP-GFP*, *aPKCAN*, *FRT*, *Crb*<sup>1A122</sup> (null) and *UAS-wgn* transgenic lines were previously published<sup>17,10,27–29</sup>, the *UAS-puc*<sup>24</sup> line

was described previously<sup>30</sup>, the *UAS-crb-intra* transgenic lines and *veli*<sup>66</sup> mutant flies were from E. Knust, the *egr*<sup>1</sup> and *egr*<sup>3</sup> mutant lines and the *UAS-ras*<sup>V12</sup>, *FRT82B*, *scrib*<sup>1</sup>/TM6b and *UAS-ras*<sup>V12</sup>, *egr*<sup>3</sup>; *FRT82B*, *scrib*<sup>1</sup>/TM6b and *yw*, *eyFlp1*; *act>y>Gal4*, *UAS-GFP*; *FRT82B*, *tubgal80* and *yw*, *eyFlp1*; *act>y>Gal4*, *UAS-GFP*, *egr*<sup>1</sup>; *FRT82B*, *tubGal80* lines were from M. Vidal, the *w*<sup>1118</sup>; +; *rnGal4*, *UAS-eiger*, *tubGal80*<sup>8</sup>/TM6B line was provided by I. Hariharan, the *yw*, *hep75/FM7*, *GFP*; *Sal*<sup>PE</sup>/cyo line was provided by F. Serras. The *rn-Gal4*, *nub-Gal4*, *puc-LacZ*, *grnd*<sup>Minos</sup> (Mi{MIC}CG10176<sup>M105292</sup>), and *grnd*<sup>Df(BSC149)</sup> lines were provided by the Bloomington *Drosophila* Stock Center.

Flies' eye pictures are acquired on a Leica Fluorescence Stereomicroscope MZ16 FA with a Leica digital camera DFC 490.

**Generation of *wgn*<sup>KO/KO</sup> mutant flies.** The *wgn* knockout allele was generated by ends-in gene targeting as described previously<sup>31</sup>. The pTV2 targeting construct carried frameshift mutations in exon 1 and 2, an I-SceI site between these two exons, a *white*<sup>+</sup> marker gene, an I-CreI site, up- and downstream homology sequences and two flanking FRT sites. The frameshift mutations were introduced by cutting the unique intrinsic restriction sites AclI and NotI followed by T4 DNA polymerase-mediated fill-in. F3 Larvae carrying the targeting construct together with an *hsp70-flp* and a *hsp70-I-SceI* transgene were heat shocked (1 h at 38 °C on day 2 and 3 after egg laying) to generate a linearized targeting construct that aligns and recombines with the endogenous wild-type locus. After crossing the F3 flies to eye-FLP-expressing flies, the F4 was screened for non-mosaic *white*<sup>+</sup> eyed flies carrying an insertion of a processed targeting construct. After processing with FLP and I-SceI, only one FRT site remained next to the *white*<sup>+</sup> marker gene in these flies. Flies still carrying the original donor insertion lost the *white*<sup>+</sup> marker gene, since it was still flanked by two FRT sites. Female F4 'w<sup>+</sup> *wgn*<sup>KO</sup>' flies, carrying two *wgn* copies on the X chromosome were crossed with males of the genotype 'v/Y; +/+; [v+, 70I-CreI] 1A/TM6'. The F5 offspring were heat shocked (1 h at 38 °C on day 2 after egg laying). F5 offspring males were selected against TM6 ('w+/–4Y; +/+; [v+, 70I-CreI] 1A/+') and typically displayed *white*<sup>+/–</sup> mosaics. When they were individually crossed to the X-chromosomal balancer stock 'bin/#852', more than 90% of the males were sterile. The F6 flies that hatched were either *white*<sup>+</sup> or *white*<sup>–</sup>. The *white*<sup>–</sup> flies were tested for the presence of the mutations. The stocks 'y w (v); P[ry+, 70FLP]4 P[v+, 70I-SceI]2B Sco/S2 Cyo' and 'v; +/+; [v+, 70I-CreI] 1A/ TM6, Ubx, e' were provided by K. Golik and Y. Rong.

**Genome-wide screen.** As an unbiased approach to identify molecules that are required for neoplastic growth, we used the *elavGal80*; *rn>avl*-RNAi/ TM6b, *tub>Gal80* testerline, which is characterized by neoplastic growth in the discs and a delay of about 2 days in pupariation. We crossed this line to the Vienna *Drosophila* RNAi Center (VDRC) collection of *phiC31* (KK) RNAi lines (10,100 lines) and screened for candidates that were able to rescue the delay at pupariation as described previously<sup>1</sup>. One-hundred and twenty-one candidates significantly rescued the pupariation delay in *rn>avl*-RNAi larvae. Of these, only 8 lines rescued the neoplastic disc growth (Extended Data Fig. 1)

**Immunostainings of larval tissues.** Tissues dissected from larvae in 1× PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.47 mM KH<sub>2</sub>PO<sub>4</sub>, pH 8) at the indicated hour AED, were fixed in 4% formaldehyde (Sigma) in PBS for 20 min at room temperature washed in PBS containing 0.1% Triton X-100 (PBT), blocked for 2 h in PBT containing 10% FBS (PBS-TF), and incubated overnight with primary antibodies at 4 °C. The next day, cells and tissues were washed, blocked in PBS-TF, and incubated with secondary antibodies at 1/500 dilution (Rhodamine Red-X-conjugated donkey anti-rat, Cy3 conjugated donkey anti-rabbit, Cy3 conjugated donkey anti-chicken, Cy3 conjugated donkey anti-mouse and Cy3 conjugated mouse anti-guinea pig, Cy5 conjugated donkey anti-rat, and Cy5 conjugated donkey anti-guinea pig from Jackson ImmunoResearch, goat anti-mouse Alexa Fluor 488, donkey anti-rat Alexa Fluor 488 and goat anti-rabbit Alexa Fluor 488 from Invitrogen) for 2 h at room temperature. After washing, tissues were mounted in Vectashield containing DAPI for staining of DNA (Vector Labs). Fluorescence images were acquired using a Leica SP5 DS (×20 and ×40 objectives) and processed using Adobe Photoshop CS5.

**Lysis and immunoprecipitations.** Cells were lysed in 100 µl buffer A (50 mM Tris-HCl (pH 8), 150 mM NaCl, 1% NP-50, 1 mM EGTA, 0.5 M NaF, phosphatase inhibitor cocktail 2 (Sigma), complete protease inhibitor cocktail (Roche)). Immunoprecipitations were performed from about 1 × 10<sup>7</sup> S2 cells. Cells were lysed in 200 µl buffer A, and cell extracts were cleared of membranous material by centrifugation at 10,000 r.p.m. for 15 min. The cleared extracts were incubated with protein G-Sepharose beads (Sigma) for 1 h to reduce unspecific binding of proteins to the beads in the subsequent purifications. Next, the precleared extracts were incubated with 80 µl of protein G-Sepharose beads and 1 µl of the relevant antibody for 3 h. Subsequently, beads were washed three times in buffer A, boiled in sample and reducing buffers (Invitrogen).



**Protein preparation.** *Drosophila* Veli full length (Veli<sup>1-195</sup>) and PDZ domain (Veli<sup>85-195</sup>) were expressed with an amino-terminal His<sub>6</sub>-tag from pET43 (Novagen) in *Escherichia coli* strain BL21 Rosetta, and purified by affinity on Ni-NTA beads (Qiagen). The Veli L27 domain (Veli<sup>1-81</sup>) was cloned in a modified pET43 plasmid as a His<sub>6</sub>-MBP-tagged construct. The fusion protein was expressed in bacteria similarly to the other constructs. After a first step of affinity purification on amylose beads (NEB), the protein was incubated with TEV protease (Invitrogen Life Technologies) to cleave the MBP tag, and loaded on a Resource-Q anion-exchange column (GE-Healthcare). The TRAF domain of *Drosophila* Traf2 (Traf2<sup>319-475</sup>) was expressed similarly, and purified by affinity followed by ion-exchange chromatography on a Resource-S column (GE-Healthcare). All Grnd constructs were expressed as GST-fusion proteins using a modified version of the pGEX-6PI vector (GE-Healthcare), and purified by affinity on glutathione sepharose (GSH) beads. The purifications were carried out at 4 °C. The PDZ domains of human Par3 (Par3-PDZ2 and Par3-PDZ3 encompassing residues 465–548 and 582–685, respectively) and *Drosophila* Bazooka (Baz-PDZ1 corresponding to residues 331–419) were expressed as GST-fusion proteins. After the first step of affinity purification on GSH-beads, the GST tag was removed from the constructs by cleavage with PreScission protease, and the PDZ domains were further purified by ion-exchange chromatography.

**In vitro binding assay.** To test the interaction properties of intracellular Grnd constructs, GST-Grnd fusion proteins (5 µM) adsorbed on GSH-beads were incubated for 40 min at 4 °C with 10 µM Veli<sup>FL</sup>, Veli<sup>PDZ</sup> or Veli<sup>PDZ</sup> in a buffer containing 10 mM HEPES pH 7.4, 30 mM NaCl, 5% glycerol, 0.5 mM EDTA and 1 mM dithiothreitol (DTT). After two washing steps, the bound species were resolved on SDS-PAGE and visualized by Coomassie staining. To test the specificity of binding, analogous pull-down assays were performed with the PDZ domains of Par3 and Bazooka under the same conditions. To monitor the ability of Traf2 to bind Grnd, 2 µM GST-Grnd<sup>125-241</sup> wild type or mutated on P200 and/or E202 were absorbed on GSH beads and incubated with 10 µM Traf2<sup>319-475</sup> in a buffer consisting of 10 mM HEPES pH 7.4, 0.15 M NaCl, 5% glycerol, 0.5 mM EDTA, 1 mM DTT and 0.05% Tween. After washing, complexes immobilized on beads were analysed by SDS-PAGE.

**Subcellular fractionation.** Subcellular fractionation of 300 wing imaginal discs dissected from wild-type animals was performed using the Qproteome Cell Compartment Kit (Qiagen) according to the manufacturer's protocol.

**Western blotting.** Proteins were resolved by SDS-PAGE using 4–12% gradient gels (NuPAGE Novex gel, Invitrogen) and transferred electrophoretically to polyvinylidene difluoride membranes (Milipore). The membranes were incubated for 1 h in blocking buffer (PBS, 5% milk) and incubated overnight at 4 °C in the same buffer containing primary antibodies at 1:1,000 dilutions: rabbit anti-Smoothed<sup>32</sup>, mouse anti-Lamin (Developmental Studies Hybridoma Bank (DSHB)), mouse anti-Dlg (4F3) from DSHB, mouse anti-Myc 9E10 (sc-40, Santa Cruz Biotechnology), rabbit anti-Myc A-14 (sc-789, Santa Cruz Biotechnology), mouse anti-Flag F3165 (Sigma), rat anti-HA 3F10 (Roche)), guinea pig anti-Grnd at 1:1,000, anti-tubulin (Sigma) at 1:2,000. The anti-Wgn antibody was used as described previously<sup>27</sup>. Membranes were washed three times in PBS-T, blocked for 1 h, and probed with secondary antibodies in blocking buffer for 1 h at room temperature. After three washes in PBS-T, chemiluminescence was observed using the ECL-Plus western blotting detection system (Amersham Biosciences). Images were generated using the Fujifilm Multi Gauge software.

**Antibodies.** To generate purified antisera specific for the Grnd protein, two peptides, CAYHPATANGKDLNA and NAINHPGSDLERAQSQ, corresponding to amino acids 140–154 and 161–176, respectively, were used as immunogens in guinea pigs (Eurogentec). The following primary antibodies were used for immunofluorescence in this study: mouse anti-Wg (4D4) from DSHB 1/50, mouse anti-Dlg (4F3) from DSHB 1/500, rat anti-E-Cad (DCAD2) from DSHB 1/500, mouse anti-MMP1 (3A6B4/5H7B11/3B8D12; the three different antibodies were mixed in amounts (1:1:1) from DSHB (1:25)), chicken anti-β-galactosidase from GeneTek 1/1,000, rabbit anti-cleaved caspase-3 (Asp 175) from Cell Signaling Technology 1/500, rabbit anti-aPKC from Santa Cruz Biotechnology 1/1,000, guinea pig anti-Grnd 1/100, rabbit anti-Veli 1/1,000 (gift from G. Pietrini), and rat anti-Crb 1/1,000 (gift from E. Knust).

**Sequence alignment and homology modelling.** Multiple sequence alignment, per cent identity matrix, and phylogenetic data were generated using Clustal Omega (<http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalo>). Data of five human death receptors of the TNFR superfamily with the highest identity with Grnd are shown. Slight manual adjustment of the alignments was performed. Amino acids are highlighted where Grnd shares identity or similarity with death receptor(s) (Fig. 1) or Grnd-like proteins from other insect species (Extended Data Fig. 2d), or where Fas shares identity or similarity with Grnd and/or Wgn (Extended Data Fig. 2b). Amino acid similarity is presented based on scores in

Gonnet PAM 250 matrix (>0.5, highly similar; ≤0.5, weakly similar). Amino acid positions are indicated in parentheses or at the end of the shown sequences. The phylogram was drawn using interactive Tree of Life<sup>33</sup> (<http://itol.embl.de/>). Number on each branch indicates branch length (Extended Data Fig. 2a). UniProt protein identifiers are: Fas (P25445), DR4 (O00220), DR5 (O14763), DcR1 (O14798), DcR2 (Q9UBN6), and Wengen (Q9VWS4). Information of protein domains in TNFR superfamily members used to construct protein domain diagram (Extended Data Fig. 2d) was obtained from UniProt (<http://www.uniprot.org/>).

**Genotypes.** Fig. 1a: w<sup>-</sup>; +; *rn-Gal4/puc-lacZ*. Fig. 1b: w<sup>-</sup>; *elav-Gal80/+; rn-Gal4, UAS avl* RNAi GD/ *puc-lacZ*. Fig. 1c: w<sup>-</sup>; *elav-Gal80/ UAS grnd* RNAi KK; *rn-Gal4, UAS avl* RNAi GD/ *puc-lacZ*. Fig. 1d: w<sup>-</sup>; *elav-Gal80/ UAS tak1* RNAi KK; *rn-Gal4, UAS avl* RNAi GD/ *puc-lacZ*. Fig. 1e: w<sup>-</sup>; *nub-Gal4/+; UAS grnd / puc-lacZ*. Fig. 1f: w<sup>-</sup>; *nub-Gal4/+; UAS grnd-intra / puc-lacZ*.

Fig. 2a: w<sup>-</sup>; *GMR-Gal4/+; +*. Fig. 2b: w<sup>-</sup>; *GMR-Gal4, UAS egr/+; MKRS/+*. Fig. 2c: wgn<sup>KO</sup>; *GMR-Gal4, UAS egr/+; MKRS/+*. Fig. 2d: w<sup>-</sup>; *GMR-Gal4, UAS egr/ UAS grnd* RNAi KK; *MKRS/+*. Fig. 2e: w<sup>-</sup>; *GMR-Gal4, UAS egr/+; MKRS/ UAS grnd-extra*. Fig. 2f: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/+; UAS egr/+*. Fig. 2g: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/ UAS tak1* RNAi KK; *UAS egr/+*. Fig. 2h–i: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/ UAS grnd* RNAi KK; *UAS egr/+*.

Fig. 3a: yw, *hsFLP; Act>CD2>Gal4, UAS GFP/+; +*. Fig. 3b: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/ UAS rh3* RNAi KK; *+*. Fig. 3c: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/ UAS grnd* RNAi KK; *+*. Fig. 3d: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/+; UAS scrib* RNAi GD/+ . Fig. 3e: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/ UAS rh3* RNAi KK; *UAS scrib* RNAi GD/+ . Fig. 3f: yw, *hs-FLP; Act>CD2>Gal4, UAS GFP/ UAS grnd* RNAi KK; *UAS scrib* RNAi GD/+ . Fig. 3g: yw, *ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; UAS ras<sup>V12</sup>+/+; FRT82B, scrib<sup>1</sup>/ FRT82B, tubGal80*. Fig. 3h: yw, *ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; UAS ras<sup>V12</sup>+/+; UAS grnd* RNAi KK; *+; FRT82B, scrib<sup>1</sup>/ FRT82B, tubGal80*. Fig. 3i: yw, *ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; egr<sup>1</sup>/ UAS ras<sup>V12</sup>, egr<sup>3</sup>; FRT82B, tubGal80/ FRT82B, scrib<sup>1</sup>*.

Fig. 4a, b: w<sup>-</sup>. Fig. 4c: w<sup>-</sup>; *UAS crb-intra/+; rn-Gal4 / +*. Fig. 4d: w<sup>-</sup>; *UAS crb-intra/ UAS puc<sup>2A</sup>; rn-Gal4 / +*. Fig. 4e: w<sup>-</sup>; *UAS crb-intra/ UAS grnd* RNAi KK; *rn-Gal4 / +*. Fig. 4f: w<sup>-</sup>; *grnd<sup>Minos</sup>/ grnd<sup>Minos</sup>; rn-Gal4 / UAS crb-intra*. Fig. 4g: w<sup>-</sup>; *UAS crb-intra/ UAS vli* RNAi; *rn-Gal4/ UAS vli* RNAi. Fig. 4h: w<sup>-</sup>; *UAS crb-intra/ nub-Gal4; vli<sup>66</sup>/ vli<sup>66</sup>*. Fig. 4i: w<sup>-</sup>; *+; UAS crbΔFBM-intra/ rn-Gal4*. Fig. 4j: w<sup>-</sup>; *UAS grnd* RNAi KK; *+; UAS crbΔFBM-intra/ rn-Gal4*. Fig. 4k: w<sup>-</sup>; *en-Gal4, UAS GFP/+; UAS aPKCΔN/ puc-lacZ*. Fig. 4l: w<sup>-</sup>; *UAS grnd* RNAi KK/ *en-Gal4, UAS GFP; UAS aPKCΔN/ puc-lacZ*.

Extended Data Fig. 1c: w<sup>-</sup>; *elav-Gal80/+; rn-Gal4, UAS avl* RNAi GD /+. Extended Data Fig. 1d: w<sup>-</sup>; *elav-Gal80/ UAS grnd* RNAi GD; *rn-Gal4, UAS avl* RNAi GD /+. Extended Data Fig. 1e: w<sup>-</sup>; *elav-Gal80/ Rh3* RNAi KK; *rn-Gal4, UAS avl* RNAi GD /+. Extended Data Fig. 1f, g: w<sup>-</sup>; *ptc-Gal4, UAS-GFP/ UAS grnd* RNAi KK; *+.*

Extended Data Fig. 4a: w<sup>-</sup>; *+; rn-Gal4 / +*. Extended Data Fig. 4b: w<sup>-</sup>; *+; rn-Gal4 / UAS grnd*. Extended Data Fig. 4c, d: w<sup>-</sup>; *+; rn-Gal4 / UAS grnd-intra*. Extended Data Fig. 4e: yw<sup>-</sup>; *sal-Gal4 / +; UAS grnd-intra / +*. Extended Data Fig. 4f: yw<sup>-</sup>; *hep<sup>75</sup>; sal-Gal4 / +; UAS grnd-intra / +*.

Extended Data Fig. 5a: w<sup>-</sup>; *GMR-Gal4, UAS egr/+; MKRS/+*. Extended Data Fig. 5b: w<sup>-</sup>; *GMR-Gal4, UAS egr/ UAS-traf2* RNAi GD; *MKRS/+*. Extended Data Fig. 5c: w<sup>-</sup>; *GMR-Gal4, UAS egr/ UAS puc<sup>2A</sup>; MKRS/+*. Extended Data Fig. 5d: w<sup>-</sup>; *GMR-Gal4, UAS egr/ UAS tak1* RNAi KK; *MKRS/+*. Extended Data Fig. 5e: w<sup>-</sup>; *GMR-Gal4, UAS egr/ UAS-wengen* RNAi IR; *MKRS/+*. Extended Data Fig. 5f: w<sup>-</sup>; *GMR-Gal4, UAS egr/ UAS-grnd* RNAi GD; *MKRS/+*. Extended Data Fig. 5g: wgn<sup>KO</sup>; *GMR-Gal4, UAS egr/ UAS-grnd* RNAi GD; *MKRS/+*. Extended Data Fig. 5h: wgn<sup>KO</sup>; *+; +*. Extended Data Fig. 5i: w<sup>-</sup>; *GMR-Gal4/ UAS-grnd* RNAi KK; *+.* Extended Data Fig. 5j: w<sup>-</sup>; *GMR-Gal4/+; UAS-grnd-extra/+*. Extended Data Fig. 5k: w<sup>-</sup>; *GMR-Gal4, UAS-ecto-egr60/+; MKRS/+*. Extended Data Fig. 5l: w<sup>-</sup>; *GMR-Gal4, UAS-ecto-egr60/ UAS puc<sup>2A</sup>; MKRS/+*.

Extended Data Fig. 6c: w<sup>-</sup>. Extended Data Fig. 6d: w<sup>-</sup>; *grnd<sup>Minos</sup>/ grnd<sup>Minos</sup>; +*. Extended Data Fig. 6e: w<sup>-</sup>; *+; rn-Gal4, UAS egr, tubGal80<sup>ts</sup> / +*. Extended Data Fig. 6f: w<sup>-</sup>; *UAS-grnd* RNAi KK/+; *rn-Gal4, UAS egr, tubGal80<sup>ts</sup> / +*. Extended Data Fig. 6g: w<sup>-</sup>; *UAS-traf2* RNAi GD /+; *rn-Gal4, UAS egr, tubGal80<sup>ts</sup> / +*. Extended Data Fig. 6h: w<sup>-</sup>; *UAS-tak1* RNAi KK/+; *rn-Gal4, UAS egr, tubGal80<sup>ts</sup> / +*. Extended Data Fig. 6i: w<sup>-</sup>; *grnd<sup>Minos</sup>/ grnd<sup>Minos</sup>; rn-Gal4, UAS egr, tubGal80<sup>ts</sup> / +*. Extended Data Fig. 6j: w<sup>-</sup>; *grnd<sup>Minos</sup>/ grnd<sup>DF</sup>; rn-Gal4, UAS egr, tubGal80<sup>ts</sup> / +*.

Extended Data Fig. 7a: yw, *ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; +; FRT82B/ FRT82B, tubGal80*. Extended Data Fig. 7b: yw, *ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; +; FRT82B, scrib<sup>1</sup>/ FRT82B, tubGal80*. Extended Data Fig. 7c: yw, *ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; UAS grnd* RNAi KK /+; *FRT82B, scrib<sup>1</sup>/ FRT82B,*



*tubGal80*. Extended Data Fig. 7d: *ey-FLP; Act>CD2>Gal4, UAS GFP/ UAS grnd* RNAi KK; +. Extended Data Fig. 7e–g: *yw, ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; UAS ras<sup>V12</sup>/+; FRT82B, scrib<sup>1</sup>/ FRT82B, tubGal80*.

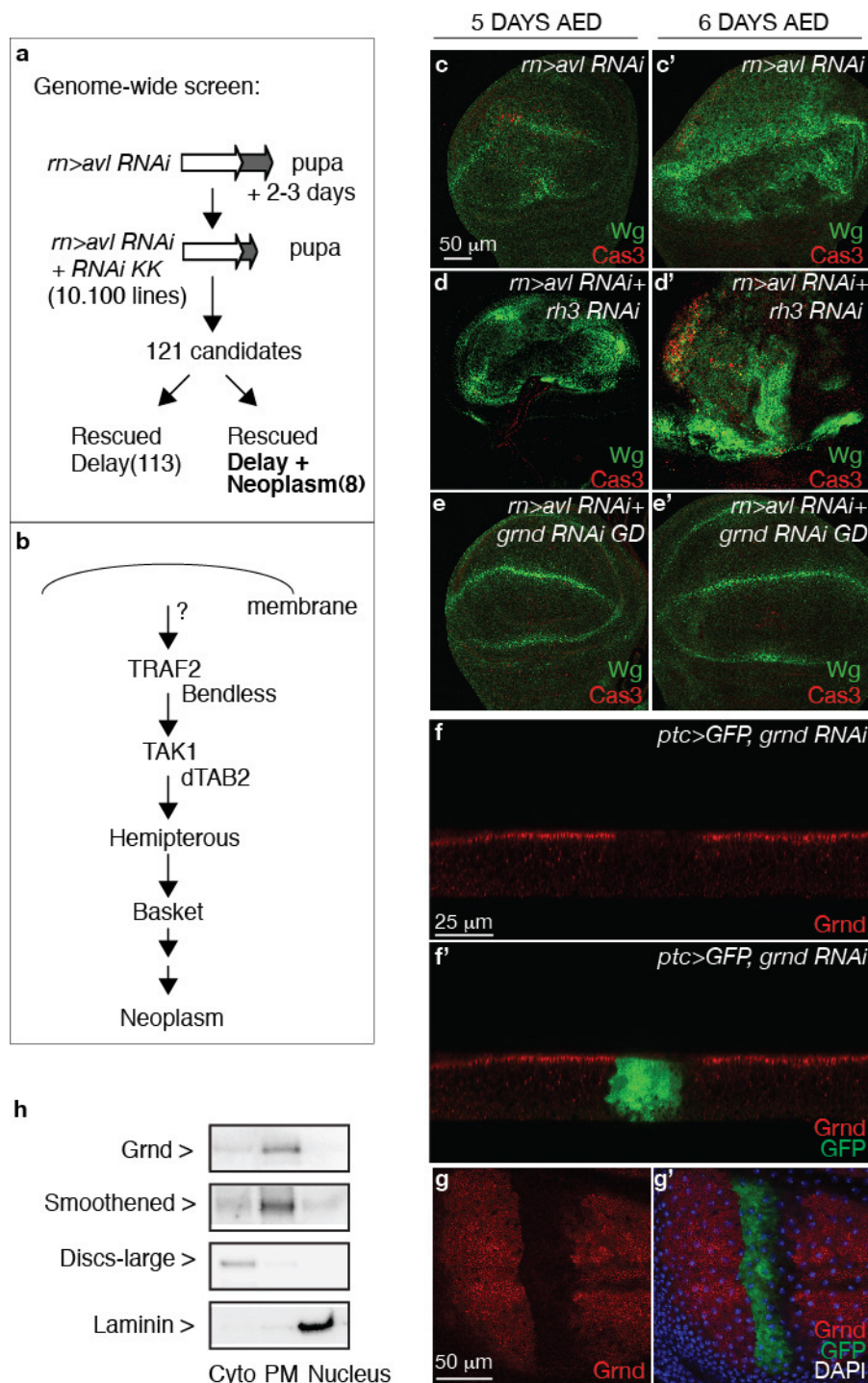
Extended Data Fig. 8a: *yw, ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; UAS ras<sup>V12</sup>/+; FRT82B, scrib<sup>1</sup>/ FRT82B, tubGal80*. Extended Data Fig. 8b: *yw, ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; UAS ras<sup>V12</sup>, UAS grnd* RNAi KK /+; *FRT82B, scrib<sup>1</sup>/ FRT82B, tubGal80*. Extended Data Fig. 8c: *yw, ey-FLP; Act>y<sup>+</sup>>Gal4, UAS GFP; UAS ras<sup>V12</sup>, UAS dilp8* RNAi KK /+; *FRT82B, scrib<sup>1</sup>/ FRT82B, tubGal80*. Extended Data Fig. 8d: *w; UAS ras<sup>V12</sup>, UAS-dlg* RNAi GD /+; *ey-FLP; Act>>Gal4, UAS GFP* /+. Extended Data Fig. 8e: *w; UAS ras<sup>V12</sup>, UAS-dlg* RNAi GD / *UAS-grnd* RNAi GD; *ey-FLP; Act>>Gal4, UAS GFP* /+. Extended Data Fig. 8f: *w; UAS ras<sup>V12</sup>, UAS-dlg* RNAi GD / *UAS-grnd* RNAi KK; *ey-FLP; Act>>Gal4, UAS GFP* /+. Extended Data Fig. 8g: *wgn<sup>KO</sup>; UAS ras<sup>V12</sup>, UAS-dlg* RNAi GD /+; *ey-FLP; Act>>Gal4, UAS GFP* /+.

Extended Data Fig. 9a: *w<sup>-</sup>; +; rn-Gal4, UAS avl* RNAi GD/+. Extended Data Fig. 9b: *w<sup>-</sup>; egr<sup>1</sup> / egr<sup>3</sup>; rn-Gal4, UAS avl* RNAi GD/+. Extended Data Fig. 9c, d: *w<sup>-</sup>; +; +*. Extended Data Fig. 9e: *w<sup>-</sup>; elav-Gal80/+; rn-Gal4, UAS avl* RNAi GD/+. Extended Data Fig. 9f: *w<sup>-</sup>; UAS crb-intra/+; rn-Gal4* /+. Extended Data Fig. 9g: *w<sup>-</sup>; elav-Gal80/+; rn-Gal4, UAS avl* RNAi GD/+. Extended Data Fig. 9h: *w<sup>-</sup>; UAS crb-intra/+; rn-Gal4* /+. Extended Data Fig. 9i: *wgn<sup>KO</sup>; UAS crb-intra/+; rn-Gal4* /+. Extended Data Fig. 9j: *w<sup>-</sup>; +; rn-Gal4/ diap-GFP*. Extended Data Fig. 9k:

*w<sup>-</sup>; UAS crb-intra/+; rn-Gal4 / diap-GFP*. Extended Data Fig. 9l: *w<sup>-</sup>; UAS crb-intra/ UAS-grnd* RNAi KK; *rn-Gal4 / diap-GFP*.

Extended Data Fig. 10f, g: *yw, hsFLP; +; FRT82B, Ubi-GFP/ FRT82B, crb<sup>1A22</sup>*. Extended Data Fig. 10h: *yw, hs-FLP; Act>CD2>Gal4, UAS GFP/ UAS veli* RNAi; *UAS veli* RNAi /+.

27. Geuking, P., Narasimamurthy, R. & Basler, K. A genetic screen targeting the tumor necrosis factor/Eiger signaling pathway: identification of *Drosophila* TAB2 as a functionally conserved component. *Genetics* **171**, 1683–1694 (2005).
28. Izaddoost, S., Nam, S. C., Bhat, M. A., Bellen, H. J. & Choi, K. W. *Drosophila* Crumbs is a positional cue in photoreceptor adherens junctions and rhabdomeres. *Nature* **416**, 178–183 (2002).
29. Zhang, L. et al. The TEAD/TEF family of transcription factor Scalloped mediates Hippo signaling in organ size control. *Dev. Cell* **14**, 377–387 (2008).
30. Martín-Blanco, E. et al. *puckered* encodes a phosphatase that mediates a feedback loop regulating JNK activity during dorsal closure in *Drosophila*. *Genes Dev.* **12**, 557–570 (1998).
31. Rong, Y. S. et al. Targeted mutagenesis by homologous recombination in *D. melanogaster*. *Genes Dev.* **16**, 1568–1581 (2002).
32. Ruel, L., Rodriguez, R., Gallet, A., Lavenant-Staccini, L. & Therond, P. P. Stability and association of Smoothened, Costal2 and Fused with Cubitus interruptus are regulated by Hedgehog. *Nature Cell Biol.* **5**, 907–913 (2003).
33. Ciccarelli, F. D. et al. Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**, 1283–1287 (2006).

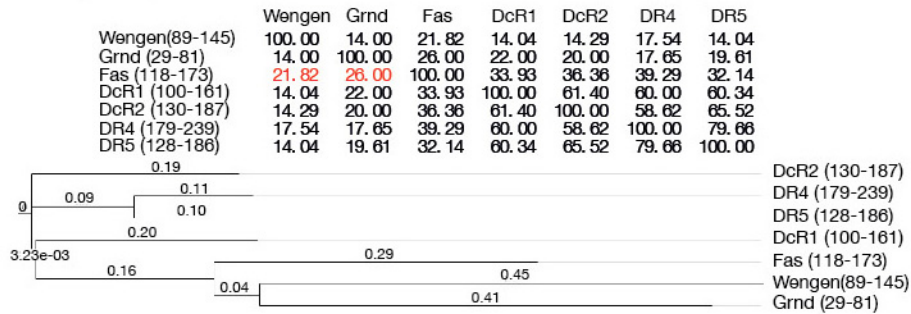
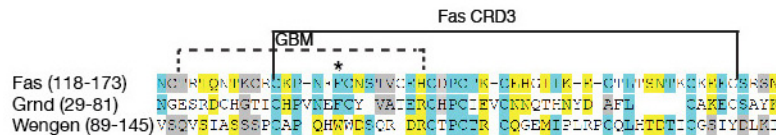
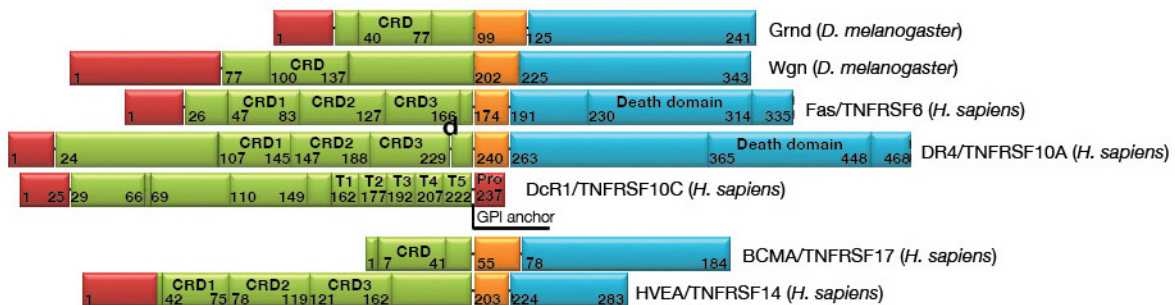


**Extended Data Figure 1 | Identification of *Grnd* in a genetic screen for molecules implicated in neoplastic growth.** **a**, Schematic of the genome-wide screen. Ten-thousand one-hundred RNAi lines were screened for their abilities to rescue the delay of the *rn>avl*-RNAi condition. Of the 121 candidates able to rescue the delay of the *rn>avl* condition, only 8 also rescued the neoplastic growth. **b**, Schematic of the five components of the JNK pathway shown to be required for the neoplastic overgrowth in the *avl* loss of function condition. Although no RNAi line targeting *Traf2* was included in the original screen, *Traf2* was subsequently shown to be required for neoplastic overgrowth in the

*avl*-RNAi condition using an alternative RNAi line. **c-e**, Wing discs of the indicated genotypes dissected 5 days (left) or 6 days (right) AED and stained for Wg (green) and Caspase3 (red). **f, g**, Transverse (**f**) or *xy* (**g**) sections of dissected wing discs expressing *grnd* RNAi in the patch domain (*GFP*, green) stained for *Grnd* (red) and DAPI (**g**, right, blue) to visualize the nuclei. **h**, Subcellular fractionation of cleared extract from dissected wing discs shows that *Grnd*, like the transmembrane receptor *Smoothed*, localizes to the membrane fraction (PM), whereas *Disc-large* remains in the cytoplasmic fraction (Cyto). Laminin serves as a control for nuclear fractions.

**a**

Percentage identity matrix

**b****c****d**

|                                |                                |                   |  |                   |     |
|--------------------------------|--------------------------------|-------------------|--|-------------------|-----|
| Drosophila simulans            | -----NGE-----                  | -----T-YKT-----   | FEPLKAEMLDIQNTQQ                           | 51                |     |
| Drosophila sechellia           | -----NGESKDCHGTI-----          | CHPVNEFCYVATERCHP | IEVCKNNQTHNYDAFLCAKE-S-YKT-----            | FEPLKAEMLDIQNTQQ  | 96  |
| Drosophila melanogaster (Grnd) | -----NGESRDCHGTI-----          | CHPVNEFCYVATERCHP | IEVCKNNQTHNYDAFLCAKE-SAYKT-----            | FEPLKAEMLDIQNTQQ  | 98  |
| Drosophila erecta              | -----NGESKDCHGTI-----          | CHPVNEFCYVATESCHP | IAVCKNNQTHNYDAFLCAKE-SAYKT-----            | FEPLKAEMLDIQNTQQ  | 98  |
| Drosophila yakuba              | -----NGESKDCHGTI-----          | CHPVNEFCYVATERCHP | IAVCKNNQTHNYDAFLCAKE-S-YKT-----            | FEPLKAEMLDIQNTQQ  | 96  |
| Drosophila ananassae           | -----NDGSQDCHGKV-----          | CHPVDEYCSAFAESCEP | APICDTTSNNYQTAMCAKE-SAFNR-----             | FEPLKAEIHNIQSNTQQ | 97  |
| Drosophila persimilis          | -----SQDGILHKDCHGKI-----       | CHPVVEYCSVFSEADS  | ATIDITKSHNYQADTCAKE-S-YNR-----             | FEPLKAEIHNIQCTQQ  | 97  |
| Drosophila willistoni          | -----TSDSNGATGKLNKGT-----      | CHPVVEYCSVFSEADS  | ATIDITKSHNYHAETSSSEIPAYNK-YETLKEEIHDIHNSNR | 102               |     |
| Drosophila virilis             | -----SNNGGDSGNSGRDCHGKL-----   | CHPVDEYCSMFAESCEP | ASICDDSSHNYHADTCAKE-SAYNS-----             | FEPLKADIHNIQSNTQQ | 105 |
| Drosophila mojavensis          | NGP-----AADSGQDCHGKV-----      | CHAVSEYCSMFAESCEP | ADIDCASSHNYHADTCAKE-SAYKS-----             | FDPLLDIQNRVVPKLL  | 102 |
| Drosophila grimshawi           | NPQSGANNA-GRDTCFGKL-----       | CHPVDEYCSMFAESCEP | SSICDANTHNYHLDITCAKE-SAYNS-----            | FEPLKAEIHNIQSNTQQ | 102 |
| Musca domestica                | ALAMTADSAATGTDCHGV-----        | CHPVVEYCSVFSEADS  | ATIDITKSHNYQADTCAKE-S-YNR-----             | FEPLKAEIHNIQCTQQ  | 103 |
| Ceratitidis capitata           | S-SSKSALAAATKDFGKV-----        | CHPVEYCSVFSEADS   | ATIDITKSHNYADVASESESVYNK-FDPLKAEIHNIHETQN  | 107               |     |
| Aedes aegypti                  | -----VALEQCHDDKPP-----         | PEDEYCDQHSYSST    | SEICSNAMTR-----VDCKKKKDYLI-----            | L-QRIATLENNLY     | 82  |
| Anopheles gambiae              | -----ATLGT-----DSAA-----       | DRYQYDESNNSRN     | SALIAKD-----E-----YSSFKKQTLRE-----         | RT-----TALESVVG   | 80  |
| Anopheles darlingi             | -----AAEGT-----ATES-----       | EQHQYQNNNNRQD     | RVLDKDK-----D-----LSSSHKPPDYLL-----        | RSLRQQVVTMQSDLW   | 79  |
| Bombyx mori                    | -----SAQITLDGIR-----GOLIQ----- | LDEYCSPEPTRNRCAP  | NVVCKNTHNYDSGLVKKEQGGYLLDLYMRSEYADPPPNL    | 88                |     |
| Danaus plexippus               | -----SSQITLDGIR-----GOLIT----- | LEEYCSPEPTRNRCAP  | SIIVCKNTHNYDSGLVKKEQGGYLLDLYMRSEYADPPPNL   | 71                |     |
| Tribolium castaneum            | -----QSEITLGETS-----GQKKK----- | DEYTSYTKAEP       | SKIKDKSTHNFDETEYKKKQDQVYHDSRYVL            | LNIEYEEL          | 90  |
| Harpegnathos saltator          | -----ADDFLPPGAK-----GNKR-----  | VSEYCSLYDTHCR     | PLDIDADHRYVLPDECEKDCQVYLHDSRYVLT           | AKQYDDL           | 91  |
| Dendroctonus ponderosae        | -----QGGITLALPK-----GQKTN----- | YHEPVSFNTQL       | LAHINQSSNNYEEVDCEKKHQNVLHDSRYVLT           | TGSTDDL           | 92  |
| Solenopsis invicta             | -----LAELDFRGT-----GEKY-----   | ITETSPFDCQHRP     | CAIACNVSSHHYQDPETKDKQLVLLDQRYAQHMDQSRQYDDL | 101               |     |
| Apis florea                    | -----TAALNHDGAK-----GEKHS----- | PIEYSSYDKSKPK     | SSINCKTARNYLPQEFSDQVYLHDSRYVL              | LNIEYEEL          | 90  |
| Apis dorsata                   | -----TAALNHDGAK-----GEKHS----- | IEYSSYDKSKPK      | SSINCKTARNYLPQEFSDQVYLHDSRYVL              | LNIEYEEL          | 89  |
| Apis mellifera                 | -----TAALNHDGAK-----GEKHS----- | IEYSSYDKSKPK      | SSINCKTARNYLPQEFSDQVYLHDSRYVL              | LNIEYEEL          | 89  |
| Bombus terrestris              | -----IAGLNPGGAD-----GEQR-----  | STVEYSSPYDKRKPK   | SSICDAGTRNYQNEIKDCQVYLHDSRYVL              | LEQYDDL           | 88  |

**Extended Data Figure 2 | Sequence homology of the extracellular domain of Grnd to CRDs of human TNFRs.** **a**, Per cent identity matrix and phylogenetic data from multiple local sequence alignment of the entire extracellular domains of Grnd and Wgn with extracellular domains (corresponding to a part of CRD2 and all of CRD3) of various human TNFR superfamily members. The extracellular domain of Grnd shares the highest identity with the GBM-harboured CRD of Fas (26%). A similar result is observed for Wgn, although it exhibits lower identity (22%) and a higher degree of divergence to Fas. **b**, Sequence alignment of Fas, Grnd and Wgn demonstrates that Grnd shares a higher degree of similarity with the GBM-harboured CRD of Fas than Wgn. This is particularly apparent in the immediate region encompassing the functionally critical F134 (marked with an asterisk) of the GBM of Fas. Blue indicates residue identity, yellow strong similarity, and grey weak similarity of Fas with Grnd and/or Wgn. **c**, Schematic representation of Grnd, Wgn, and

various members of the human TNFR superfamily showing a high level of diversity in the number and arrangement of CRDs. The colour codes used are: red for signal peptide, green for extracellular domains, orange for transmembrane domains, and blue for intracellular domain. Numbers in each box indicate corresponding amino acid positions. Pro, prodomain; T1–5, TAPE1–5 (threonine-, alanine-, proline- and glutamine-rich repeats). **d**, Multiple sequence alignment showing the conservation of the extracellular domain of Grnd in other insect species. Whereas global similarity of the extracellular domain of Grnd degenerates outside the melanogaster group, the cysteine residues of the CRD (highlighted in red) and the aromatic residue (F/Y, marked with an asterisk) corresponding to F134 of Fas CRD3 remain highly conserved throughout the Hymenoptera order. Blue indicates residue identity; yellow, strong similarity; and grey, weak similarity to Grnd.

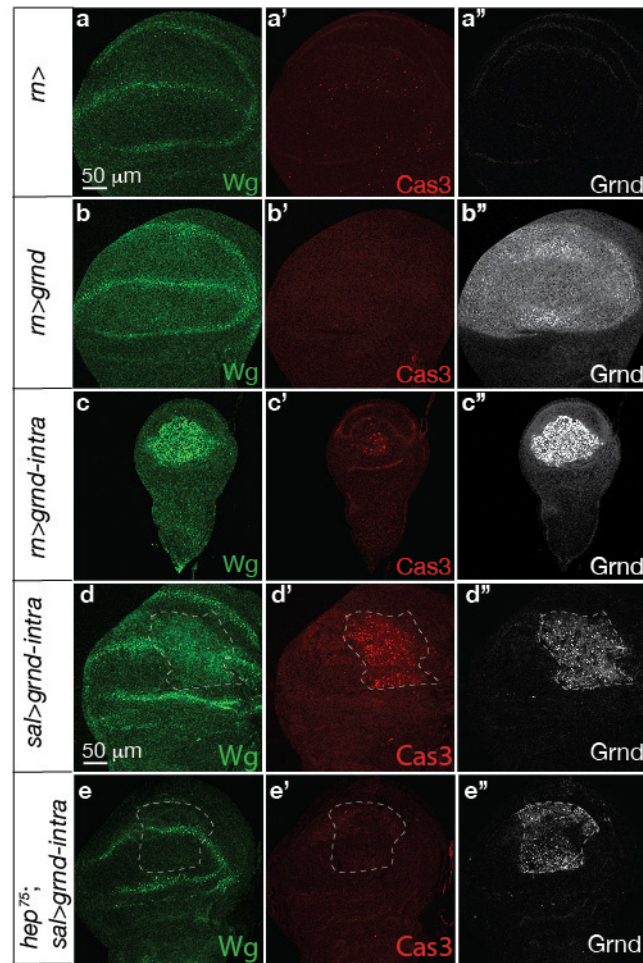


|                                       |  |                                       |                      |                                 |                  |          |
|---------------------------------------|--|---------------------------------------|----------------------|---------------------------------|------------------|----------|
| <i>Drosophila simulans</i>            | GNRCFQKL-LRRL                                | QSKAYPHPATANGKDLN-AT                  | TIQN                 | LNAINHPGSDLE                    | RAQSQIYSVA       | 130      |
| <i>Drosophila sechellia</i>           | GNRCFQKL-LRRL                                | QSKAYPHPATANGKDLN-AT                  | TIQN                 | LNAINHPGSDLE                    | RAQSQIYSVA       | 175      |
| <i>Drosophila melanogaster (Grnd)</i> | GNRCFQKL-MRRL                                | QSKAYPHPATANGKDLN-AT                  | TIQN                 | LNAINHPGSDLE                    | RAQSQIYSVA       | 177      |
| <i>Drosophila erecta</i>              | SNRCFQKL-LRRL                                | QSKAYPHPATANGKDLN-AT                  | TIQN                 | LNAINHPGSDLE                    | RAQSQIYSVA       | 177      |
| <i>Drosophila yakuba</i>              | SNRCFQKL-LRRL                                | QSKAYPHPATANGKDLN-AT                  | TIQN                 | LNAINHPGSDLE                    | RAQSQIYSVA       | 175      |
| <i>Drosophila ananassae</i>           | HKRCIQRM-LQRM                                | QSKAYP-PGTANGKDLN-AT                  | TIQN                 | LNAINRHGSDLE                    | RAPSQIYSVA       | 174      |
| <i>Drosophila persimilis</i>          | RKRCVQQM-LQRY                                | QSKAYPHPGTANGKDLN-AT                  | TIQN                 | LSAI-RHGSDIE                    | RAPSQIYSVA       | 174      |
| <i>Drosophila willistoni</i>          | RKRCVQQM-LQKL                                | QSKSYPHQTNVNGKDLN-AT                  | TIQN                 | LNAINRHGSDIE                    | RAPSQIYSVA       | 182      |
| <i>Drosophila virilis</i>             | HSSCVQRL-KSRL                                | LPK-PQSNANGKDLN-AT                    | TMQN                 | PNAFYR-DIE                      | RAPSQINSVA       | 178      |
| <i>Drosophila mojavensis</i>          | VKRCLORC-MARL                                | QSKTOP-SNGKELN-GM                     | TIQN                 | PNAFNR-DIE                      | RAPSQIYSVA       | 178      |
| <i>Drosophila grimshawi</i>           | HKRCMROL-LSRL                                | VSK-PHOSNANGKDLN-AT                   | TIQN                 | LNAINR-DIE                      | RAPSQIYSVA       | 163      |
| <i>Musca domestica</i>                | HKRYVQHA-LKKL                                | KLKKDPLPPTAHNGGGGGT                   | TIQN                 | MCVIN-DIE                       | RSPSQIYSMT       | 158      |
| <i>Ceratitis capitata</i>             | RKRFTAM-LKKL                                 | HKQ-HEAAGTGKDLVVGGA                   | TIPN                 | VMAING-SNSTALPPLPQVATTAPPSTYM   |                  | 160      |
| <i>Aedes aegypti</i>                  | KSNHFSWSFLKRL-LAREKSNPPT                     | MEYTHENPHT                            | KSPK                 | L-STRSGSNSTALPPLPQVATTAPPSTYM   |                  | 145      |
| <i>Anopheles gambiae</i>              | AIREF-CRLPQL                                 | NKKEQGP-VGYTHENPNT                    | KSPK                 | A-VPRNGANAAR                    | PPAS-TVSTYTP-E   | 128      |
| <i>Anopheles darlingi</i>             | SVDFE-SRCPQOGKPSSSAGAP                       | VGYTHENPNT                            | KSPK                 | L-AVKINGAAGIPKQQQVAPS-VALSYTP-E |                  | 134      |
| <i>Bombyx mori</i>                    | LKOKFQPAKINVKQ-Y                             | PNDLTHHN                              |                      | PHAEAPKPKPELKLE                 | IRNPDPRLRNQPLNVK | 157      |
| <i>Danaus plexippus</i>               | IMKKLQPSKNRVKH-Y                             | PTDLTHHN                              |                      | PHAEAPKPKPELKLE                 | IRNPDPRLRNQPLNVK | 1346     |
| <i>Tribolium castaneum</i>            | KKNITLATIKNKLFSSKADSA                        | QNKSTPVS                              |                      | NOVKNEKPDRLS                    | ISPES            | 163      |
| <i>Harpegnathos saltator</i>          | IQNTLRLFRKNWAKKAANKN                         |                                       | KVQDDVEANATKONGKLKLT | MPTISATVDNRS                    | STG              | 171      |
| <i>Dendroctonus ponderosae</i>        | KKNITLAVIKGMFGKNDSSA                         | TTNST                                 |                      | SNQDNKKRDLRL                    | IPSPT            | 151      |
| <i>Solenopsis invicta</i>             | IQKSLQSVFRNKKV-KANKN                         |                                       | KVQDDVEAAVNKONGKLKLT | IPNISATVEQESKIENNNS             |                  | 174      |
| <i>Apis florea</i>                    | MRAAVGRITFASKKIKRMSNN                        | NDNNVGRIN                             |                      |                                 |                  | 160      |
| <i>Apis dorsata</i>                   | MRAAIRGSFASKKIKRMSNN                         | NDNNVSKNNNNNDNRNATLQDAELGTSKONGKLKLT  |                      | MPTISASVAPSRYSENE               |                  | 188      |
| <i>Apis mellifera</i>                 | MRAAIRRSFASKKIKRMSNN                         | SDNSVGNNDNNNDNRNATLQDAELGTSKONGKLKLT  |                      | MPTISASVAPSRYSENE               |                  | 186      |
| <i>Bombus terrestris</i>              | MQTTLRAIAGKRAKGTNGS                          | GGDNDNNNDNRGKSTLRDAESGTSRHNGKLKLT     |                      | MPTISASVAPSRILENE               |                  | 184      |
|                                       |  |                                       |                      |                                 |                  |          |
| <i>Drosophila simulans</i>            |  | GAAEGSVVIMTTPVSTRYPAENSTTPTTVMTEI     |                      | GYGYDNQAMVTPVSEKPSAATI          | PVAF             | 194      |
| <i>Drosophila sechellia</i>           |  | GAAEGSVVIMTTPVSTRYPAENSTTPTTVMTEI     |                      | GYGYDNQAMVTPVSEKPSAATI          | PVAF             | 239      |
| <i>Drosophila melanogaster (Grnd)</i> |  | GAAEGSVVIMTTPVSTRYPAENSTTPTTVMTEI     |                      | GYGYDNQAMVTPVSEKPSAATI          | PVAF             | 241      |
| <i>Drosophila erecta</i>              |  | GAAEGSVVIMTTPVSTRYPAENSTTPTTVMTEI     |                      | GYGYDNQAMVTPVSEKPSAATV          | AAAF             | 241      |
| <i>Drosophila yakuba</i>              |  | GAAEGSVVIMTTPVSTRYPAENSTTPTTVMTEI     |                      | GYGYDNQAMVTPVSEKPSAATV          | AAAF             | 239      |
| <i>Drosophila ananassae</i>           |  | GAAEGSVLIMTTPVSTRYPAENSTTPTTVTEV      |                      | EYGYDNQAMVTPVSEKPT              | SVA-AAAF         | 238      |
| <i>Drosophila persimilis</i>          |  | GAAEGSVLILITTPVSTRYPAENSTTPTTVTEI     |                      | GYGYDNQAMVTPVSEKPSAAT           | IAAF             | 238      |
| <i>Drosophila willistoni</i>          |  | GPADGSVVIMTTPVSTRYPAENSTTPTTVTEI      |                      | GYGYDNQAMVTPVSEKPSV-TATTGPAP    |                  | 246      |
| <i>Drosophila virilis</i>             |  | GAGEGSVLITATPVSTRYPAENSTTPTTVTEI      |                      | GYVYDNQAMVTPVTEKPGP-TA          | MAAF             | 242      |
| <i>Drosophila mojavensis</i>          |  | GAAEGSVLIMTTPVSTRYPAENSTTPTTVTEI      |                      | GYGYDNQAMVTPVSEKPSAATV          | AAAF             | 242      |
| <i>Drosophila grimshawi</i>           |  | GAAEGSVLITATPVSTRYPAENSTTPTTVTEI      |                      | GYGYDNQAMVTPVSEKPSAATV          | AAAF             | 240      |
| <i>Musca domestica</i>                |  | G-VEGSLVIMTTPISRRHAEQTPSPQAVT         |                      | NEYSYDNQALAVTPVSEKPGGG          | SVVE             | 240      |
| <i>Ceratitis capitata</i>             |  | GSDVGSVQIVTTPINGRYPAENSTTPTTV         |                      | SGEYSYDNQALAVTPVSEKNGS          | VHGF             | 240      |
| <i>Aedes aegypti</i>                  |  | GSDVGSVQIVTTPINGRYPAEDWRE             |                      | NHAYDNGGLVVTTP                  | TENTYEVVEAKMMNR  | 227      |
| <i>Anopheles gambiae</i>              |  | TEADNSVQGTITTSIRHYPAEDSTE             |                      | SYSDNACNVTP                     | TSQNP            | MPKY 209 |
| <i>Anopheles darlingi</i>             |  | TEAETSVOGTITTSIRHYPAEDSTE             |                      | SYSDNACNVTP                     | TSVAP            | MPKY 216 |
| <i>Bombyx mori</i>                    |  | DLETRTSQSDKSQGATTPKTIISTALSRRHAEEDTTL |                      | DFSYDNMGNNVTPPGQPAAS            | HKF              | 239      |
| <i>Danaus plexippus</i>               |  | DLDART-QTDKSQGATTPKTIISTALSRRHAEEDTTL |                      | DFSYDNMGNNVTPPEQAT-S            | HRF              | 1421     |
| <i>Tribolium castaneum</i>            |  | THSGHSPVTITTSISRRHAEEDTTL             |                      | DYAYDNPMAMSSPNKNDPN             | NSF              | 227      |
| <i>Harpegnathos saltator</i>          |  | NNNDTPETITTSLSRRHAEEDTTL              |                      | DYAYDNPMAMTSPPEAQLRTKRE         | SSF              | 234      |
| <i>Dendroctonus ponderosae</i>        |  | VNSDHSPTVITTSISDRRPAEDSTL             |                      | DYAYDNPMAMSSNNISY               |                  | 215      |
| <i>Solenopsis invicta</i>             |  | NSNSNSNGINTIPNITSTLSRRHAEEDTTL        |                      | DYAYDNPMAMTSPDSSA               |                  | 244      |
| <i>Apis florea</i>                    |  |                                       |                      |                                 |                  | 160      |
| <i>Apis dorsata</i>                   | TRHG-NGSSAGSGSGSGTAVSNVTPNITSTLSRRHAEEDTTL   |                                       |                      | DYAYDNPMAMTSPESVQLRTKRE         | SSF              | 272      |
| <i>Apis mellifera</i>                 | TRHG-NGSSAGSGSGSGTAVSNVTPNITSTLSRRHAEEDTTL   |                                       |                      | DYAYDNPMAMTSPESVQLRTKRE         | SSF              | 270      |
| <i>Bombus terrestris</i>              | SGHGGNGTGSASGSGSANGSGTGNATPNITSTLSRRHAEEDTTL |                                       |                      | DYAYDNPMAMTSPETVQRLTKRE         | SSF              | 268      |

## TRAF6-like binding motif

**Extended Data Figure 3 | Identification of a conserved TRAF6-like binding motif in the intracellular domain of Grnd proteins.** Multiple sequence alignment of the intracellular domain of Grnd and Grnd-like proteins in other species. A binding motif for TRAF6 (PxExx(aromatic/acidic residue)), the closest homologue of *Drosophila* Traf2, was previously identified in human TNFRs<sup>4</sup>. The position of a related putative Traf2-binding motif in the

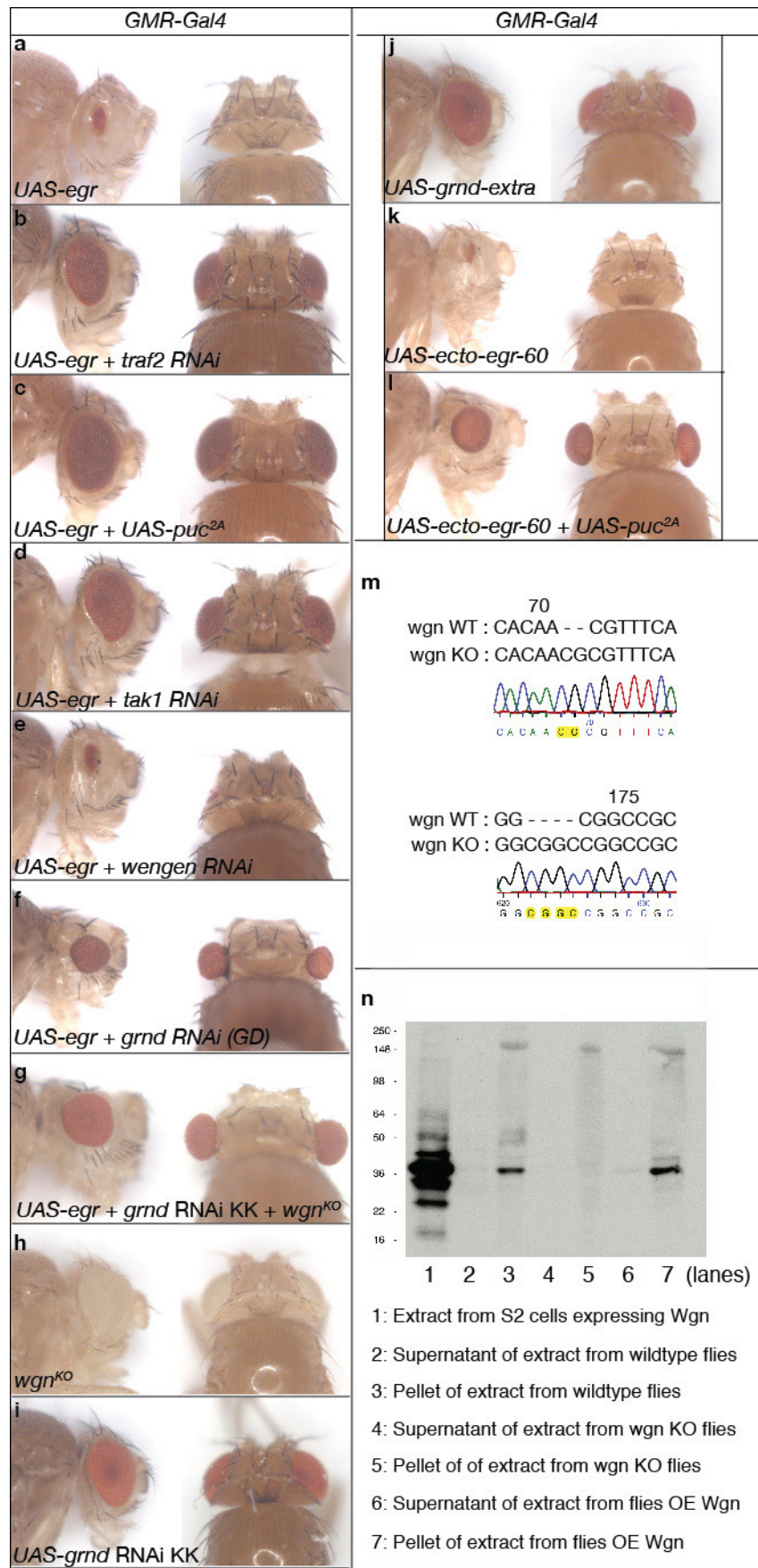
intracellular domain of Grnd is indicated. Note that, although sequence similarity of the intracellular domain of Grnd degenerates outside the family *Drosophilidae*, the core residues of this motif (highlighted in red) remain highly conserved throughout the order Hymenoptera. Blue indicates residue identity; yellow, strong similarity to Grnd.



**Extended Data Figure 4 | Grnd acts upstream of JNK signalling.** **a–d**, Wild-type wing discs (**a**) or discs expressing *grnd* (**b**) or *grnd-intra* (**c**, **d**) in the *rn* domain were dissected 5 days AED and stained for Wg (left; green), Cas3 (middle; red), and Grnd (right; grey). **d**, **e**, Wing discs expressing *grnd-intra* in

the *sal* domain (marked by a truncated line) in a wild-type (**d**) or *hep*<sup>75</sup> mutant (**e**) background were dissected 5 days AED and stained for Wg (left; green), Cas3 (middle; red), and Grnd (right; grey).



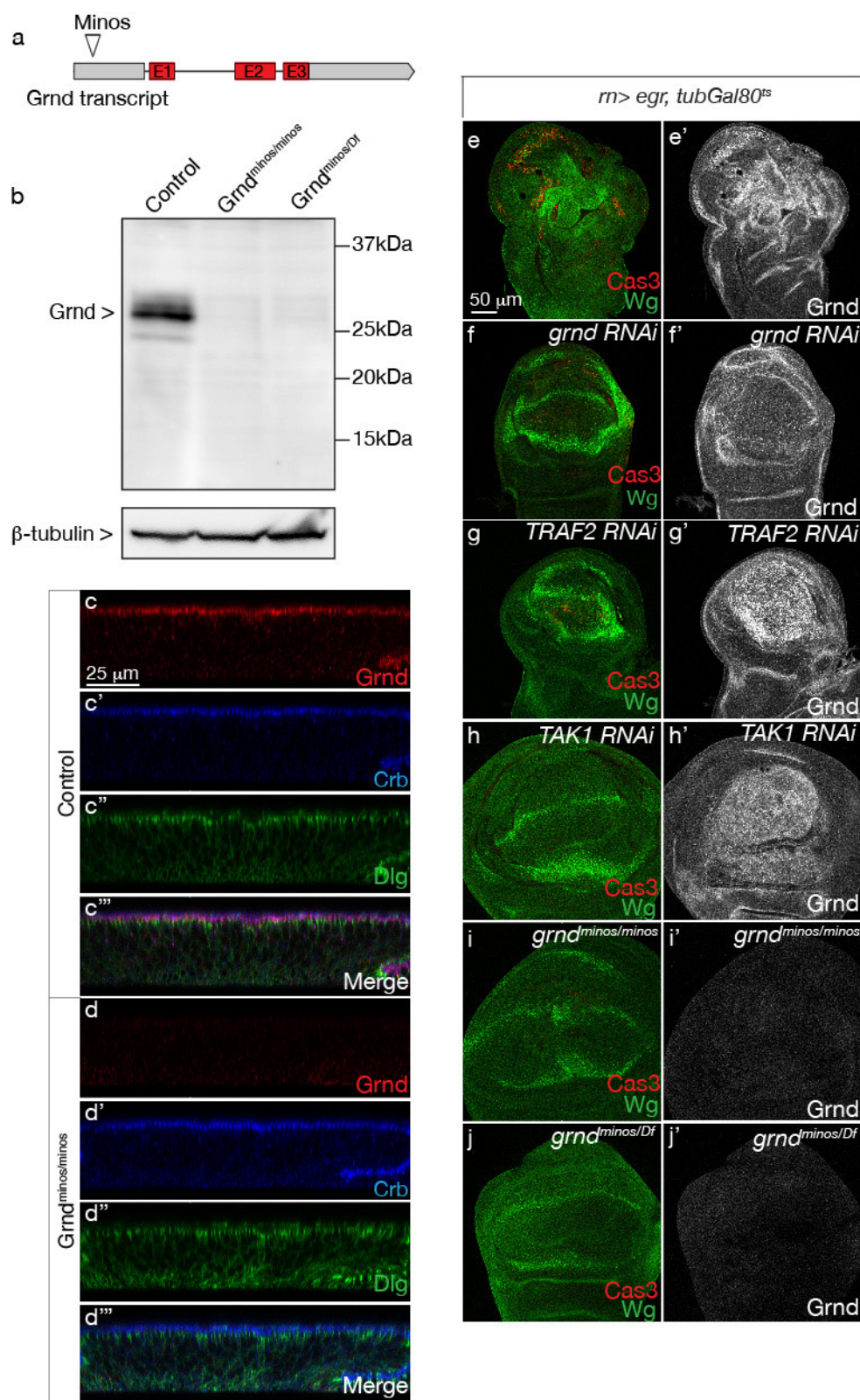


**Extended Data Figure 5 | Characterization of the *wgn*<sup>KO</sup> mutant allele.**

**a–e**, Light micrographs of *Drosophila* adult eyes are shown. Original magnification,  $\times 30$ . The small-eye phenotype caused by *GMR*-driven expression of *egr* (**a**) is rescued by co-expressing *puc* (**c**) or reducing the activity of Traf2 (**b**), but not by reducing Wgn activity (**d**; also serves as a control to exclude a Gal4 titration effect). **f, g**, Reducing Grnd activity partially rescues the small-eye phenotype, giving rise to the hanging-eye phenotype (**f**) and is not further rescued in a *wgn*<sup>KO</sup> mutant background (**g**). **h**, *wgn*<sup>KO</sup> mutant flies do not display any eye phenotype. **i, j**, Reducing Grnd levels or expressing Grnd-extra in the *GMR* domain does not affect eye morphology. **k**, *GMR*-Gal4-

induced expression of a soluble form of Egr lacking the transmembrane and cytoplasmic domain (Ecto-Egr-60) gives rise to a small-eye phenotype<sup>10</sup>. **l**, Co-expression of *puc*<sup>2A</sup> in a *GMR>ecto-egr-60* background partially rescues the small-eye phenotype, giving rise to a hanging-eye phenotype<sup>10</sup>. **m**, Sequence analysis of DNA extracted from flies homozygous for the *wgn*<sup>KO</sup> mutation showing the two frameshift mutations introduced by cutting the unique intrinsic restriction sites AclI (exon 1) and NotI (exon 2) followed by T4 DNA polymerase-mediated fill-in. **n**, Western blot of fly extracts from the indicated genotypes probed with an anti-Wgn antibody<sup>27</sup>. KO, knockout; OE, overexpressing; WT, wild type.

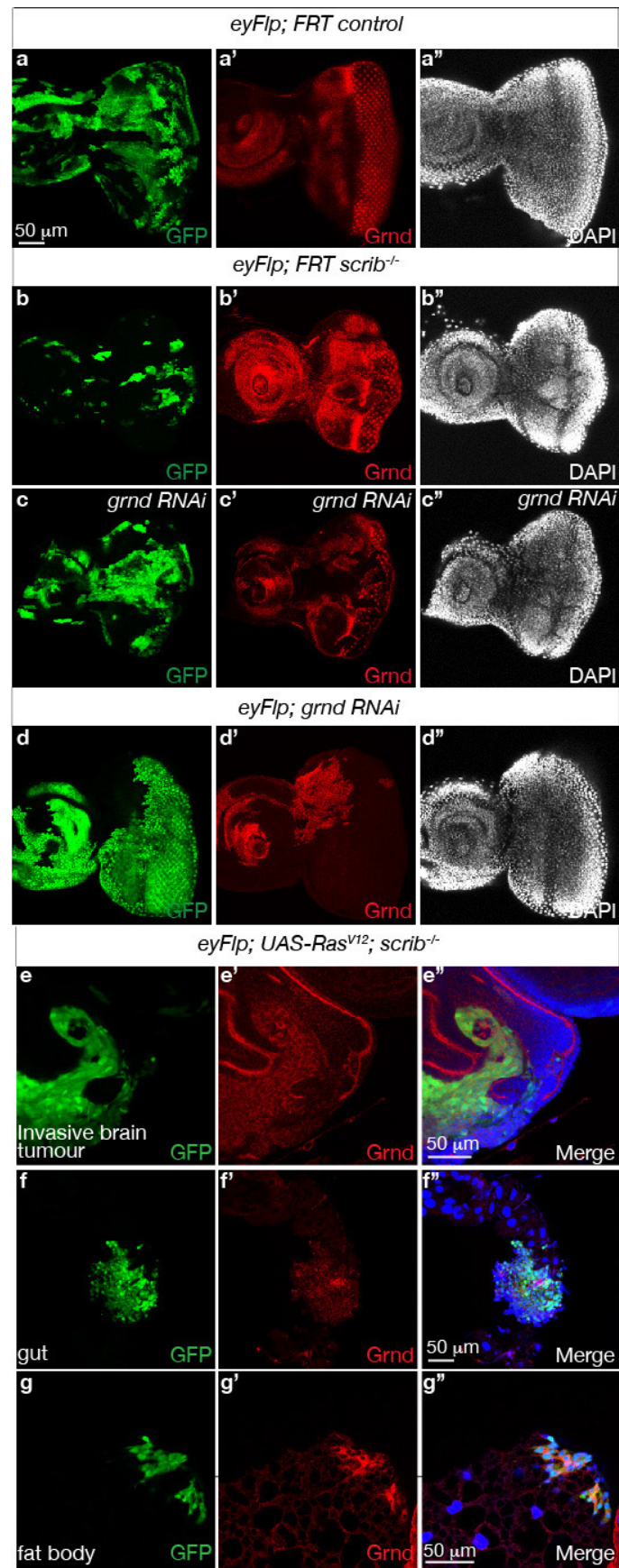




#### Extended Data Figure 6 | Characterization of the *grnd*<sup>Minos</sup> mutant.

**a**, Schematic representation of the Minos transposon insertion in the *grnd* gene region corresponding to the 5' untranslated region of the transcript. **b**, Western blot of wing disc extracts from the indicated genotypes probed with anti-Grnd (top) and anti- $\beta$ -tubulin (bottom) antibodies. Grnd protein migrates at the predicted size of 27 kDa (lane 1, control). Grnd protein levels are decreased below detection level in both *grnd*<sup>Minos/Minos</sup> and *grnd*<sup>Minos/Df</sup> mutant flies, indicating that *grnd*<sup>Minos/Minos</sup> is indistinguishable from a null mutant. **c, d**, Polarity is not affected in *grnd*<sup>Minos/Minos</sup> and *grnd*<sup>Minos/Df</sup> wing discs.

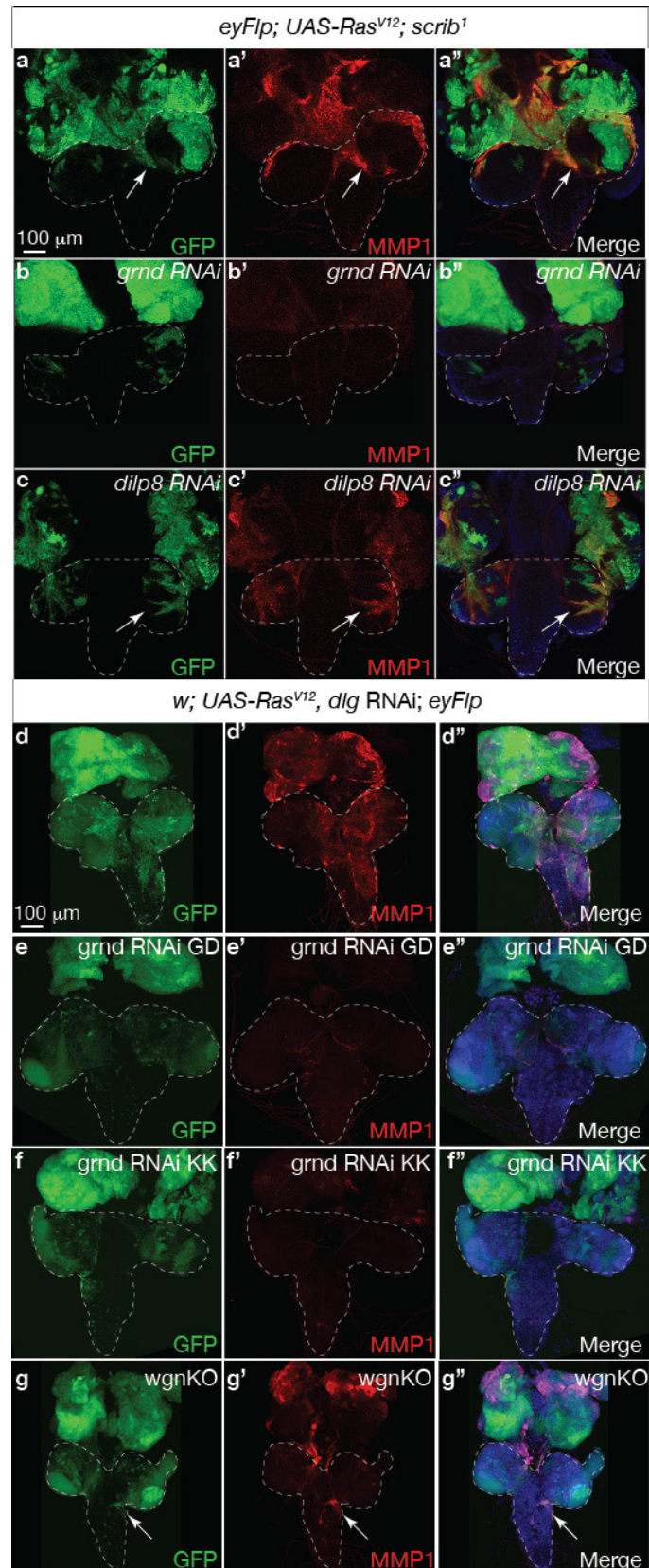
Transverse sections of dissected control (c) or *grnd*<sup>Minos/Minos</sup> (d) wing discs stained for Grnd (top; red), Crb (middle top; blue) and Dlg (middle bottom; green). **e–j**, Wing discs dissected from third instar larvae after 40 h of Egr expression in the *rn* domain. Pouch ablation, ubiquitous Wg expression and apoptosis observed in *rn>egr, Tub-Gal80<sup>ts</sup>* discs (e) are suppressed to various extents upon co-expression of *grnd*-RNAi (f) or reduction of JNK activity (g, h) and fully suppressed in *grnd*<sup>Minos/Minos</sup> (i) and *grnd*<sup>Minos/Df</sup> (j) mutant backgrounds.



**Extended Data Figure 7 | Grnd is highly expressed in  $Ras^{V12}/scrib^{-/-}$  micro-metastases.** **a–c**, Reducing *grnd* levels in *scrib*<sup>-/-</sup> mutant clones prevent their elimination. Eye discs dissected 5 days AED carrying GFP-labelled MARCM clones of the indicated genotypes stained for Grnd (middle; red) and DAPI (right; grey). **d**, Reducing Grnd levels does not affect cell viability. Eye discs dissected 5 days AED carrying GFP-labelled *grnd* RNAi Flip-

out clones stained for Grnd (middle; red) and DAPI (right; grey). **e–g**, Invasive brain tumours and micro-metastases display high levels of Grnd protein. Dissected eye–brain complex (**e**), gut (**f**) and fat body (**g**) displaying GFP-labelled invasive  $Ras^{V12}/scrib^{-/-}$  clones (**e**; green) or micro-metastases (**f**, **g**; green) stained for Grnd (red) and DAPI (blue).

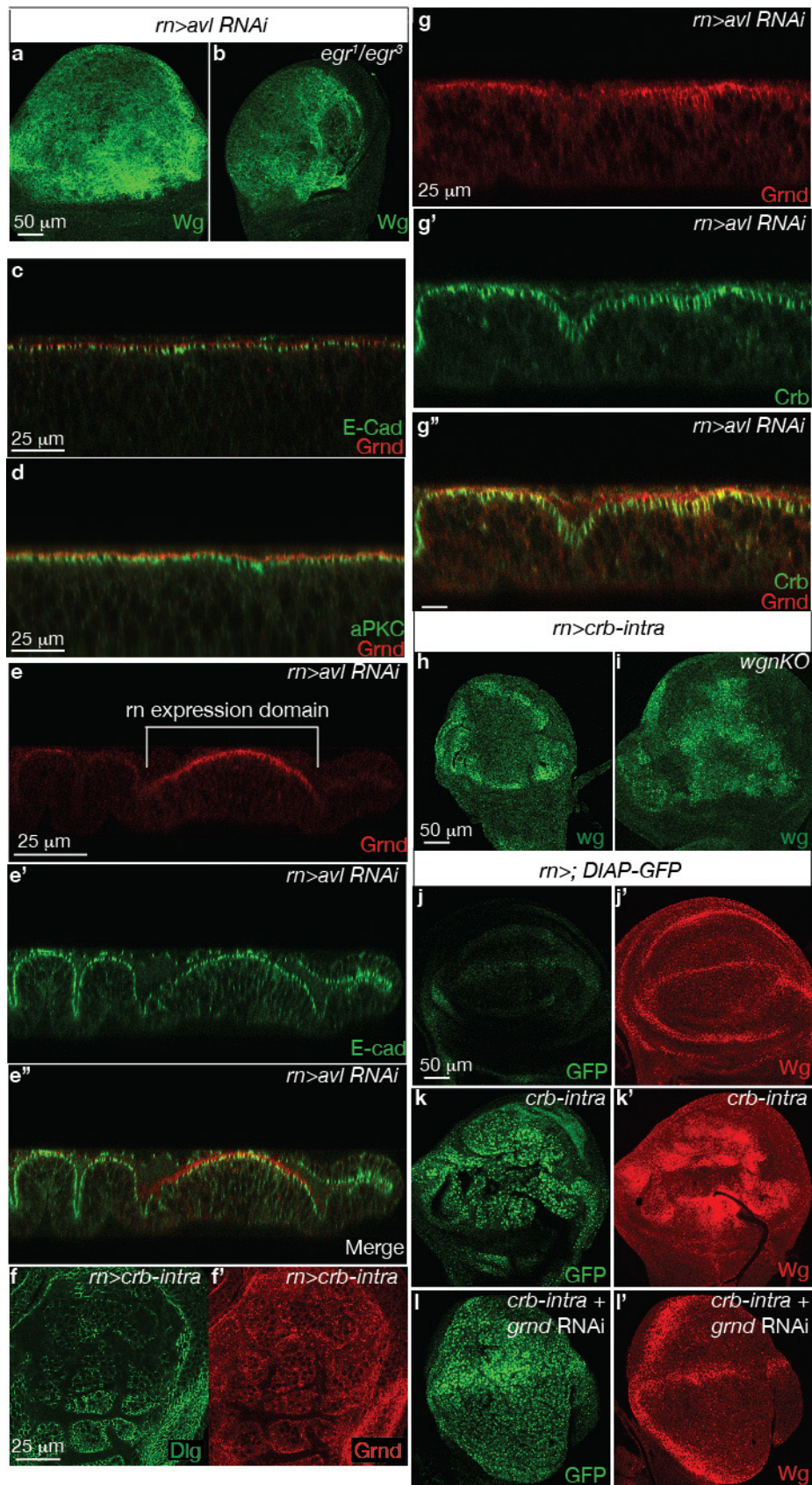




**Extended Data Figure 8 | Grnd, but not Wgn, is required for Mmp1 expression both in  $Ras^{V12}/scrib^1$  and  $Ras^{V12}/Dlg$  RNAi tumours.** **a–c,** Eye-brain complexes dissected 7 days AED carrying  $Ras^{V12}/scrib^{-/-}$  (**a**),  $Ras^{V12}/scrib^{-/-} + grnd$ -RNAi (**b**), or  $Ras^{V12}/scrib^{-/-} + dilp8$ -RNAi (**c**; control RNAi to exclude a Gal4 titration effect) clones labelled by GFP and stained for Mmp1 (middle; red).  $Ras^{V12}/scrib^{-/-}$  and  $Ras^{V12}/scrib^{-/-} + dilp8$ -RNAi clones

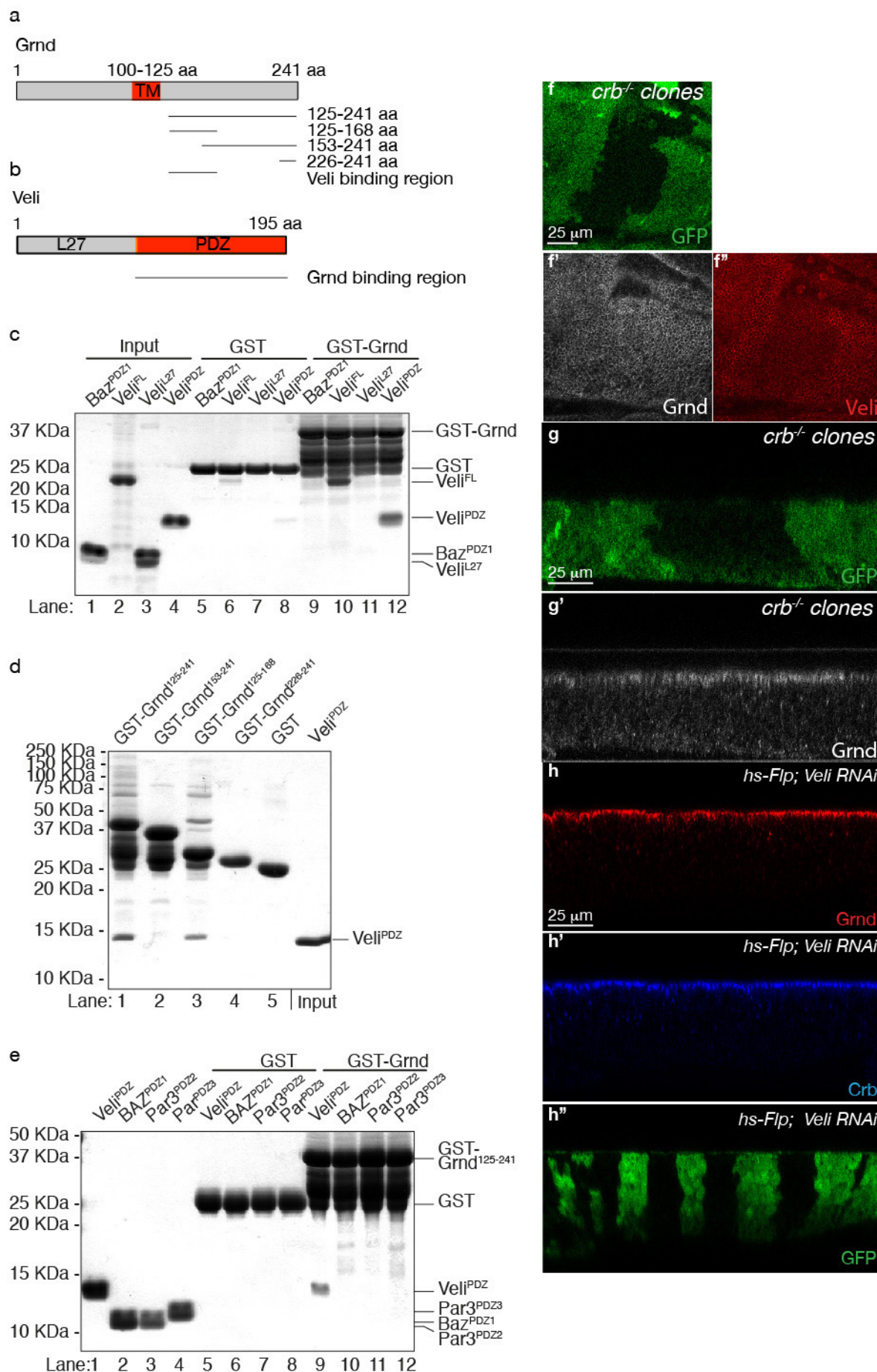
invade the ventral nerve cord (**a, c**; see white arrow). **d–g,** Eye-brain complexes dissected 7 days AED carrying  $Ras^{V12}/Dlg$ -RNAi (**d**) or  $Ras^{V12}/Dlg$ -RNAi +  $grnd$ -RNAi (**e, f**) clones (labelled by GFP), or  $Ras^{V12}/Dlg$ -RNAi GFP-labelled clones in a  $wgn$ -null mutant background (**g**) stained for Mmp1 (middle; red) and DAPI (right; blue).  $Ras^{V12}/Dlg$  RNAi clones invade the ventral nerve cord in  $wgn$ -null mutant animals (**g**; see white arrow).





**Extended Data Figure 9 | Egr is not required for the *rn>avl* RNAi-induced neoplastic growth.** **a, b**, Wing discs of the indicated genotypes dissected 6 days AED and stained for Wg (green). **c, d**, Transverse sections of dissected wild-type discs stained for Grnd (red) and either E-cad (**c**; in green), or aPKC (**d**; in green). **e, g**, Transverse sections of *rn>avl* RNAi wing discs dissected 5 days AED and stained for Grnd (top, bottom; red) and either E-cad (**e**, middle, bottom; green) or Crb (**g**, middle, bottom; green). The *rn* expression domain is indicated (**e**, top). **f**, Wing discs expressing *crb-intra* in the *rn* domain (*rn>crb-intra*) dissected 6 days AED and stained for Disc-large (Dlg, green) to visualize

neoplastic structures and Grnd (red). **h, i**, Wing discs expressing *crb-intra* in the *rn* domain of wild-type (**h**) or *wgn*-knockout mutant (**i**) animals dissected 6 days AED and stained for Wg (green). **j–l**, Reducing *grnd* levels does not affect *crb-intra*-mediated inhibition of Hippo signalling, which causes an upregulation of the Yki transcriptional target Diap1. Control discs (**j**) or wing discs expressing Crb-*intra* alone (**k**) or together with *grnd* RNAi (**l**) in the *rn* domain were dissected from animals carrying a reporter gene for Hippo pathway activity (*Diap-GFP*, green) and stained for Wg (red).



**Extended Data Figure 10 | Grnd directly interacts with the PDZ domain of Veli and neither depends on Veli nor Crb for its proper localization.** **a, b,** Schematic representation of Grnd and Veli (also known as Lin-7) domains and the various truncations used to map the interface of the Grnd–Veli interaction. L27, Lin-2/Lin-7 domain; PDZ, PSD-95/Dlg/ZO-1 domain; TM, transmembrane domain. **c,** GST–Grnd<sup>125–241</sup> immobilized on GSH beads was incubated with full-length Veli (Veli<sup>FL</sup>) or the complementary Veli fragments L27 (Veli<sup>L27</sup>) and PDZ domain (Veli<sup>PDZ</sup>). The PDZ1 domain of the polarity protein Bazooka was included in the pull-down as a specificity control (Baz<sup>PDZ1</sup>). Coomassie-blue stained SDS–polyacrylamide gel electrophoresis (SDS–PAGE) was used to visualize species retained on beads. Veli<sup>FL</sup> and Veli<sup>PDZ</sup> bind to Grnd (lanes 10 and 12), but not Veli<sup>L27</sup> nor Baz<sup>PDZ1</sup> (lanes 11 and 9), indicating that Grnd specifically recognizes the PDZ domain of Veli. **d,** To map the interface between Grnd and Veli<sup>PDZ</sup>, the same binding assay was

performed using a battery of Grnd deletion mutants adsorbed on GSH beads. Veli<sup>PDZ</sup> forms a complex with Grnd<sup>125–241</sup> and Grnd<sup>125–168</sup> (lanes 1 and 3), but not with the carboxy-terminal region of Grnd (lanes 2 and 4) or GST alone (lane 5, control). **e,** To address further the specificity of the binding of Grnd<sup>125–241</sup> to the PDZ domain of Veli, we repeated the pull-down assay with the PDZ domains of Baz and its human orthologue PAR3. Veli<sup>PDZ</sup> is retained on GST–Grnd<sup>125–241</sup> beads (lane 9), while PAR3<sup>PDZ2</sup>, PAR3<sup>PDZ3</sup> and Baz<sup>PDZ1</sup> are not (lanes 10, 11 and 12). **f–g,** Reducing Crb levels affects the localization of Veli, but not Grnd. **f–g, xy (f)** or transverse (**g**) sections of dissected wing discs bearing *crb*<sup>−/−</sup> mutant clones (labelled by absence of GFP) stained for Grnd (**f**, bottom left; **g**, bottom; white) or Veli (**f**, bottom right; red). **h,** Reducing Veli levels does not affect Crb<sup>20</sup> or Grnd localization. Transverse sections of dissected wing discs bearing clones expressing *veli* RNAi (**h**, bottom; labelled by the presence of GFP) stained for Grnd (top; red) and Crb (middle; blue).



# Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117

Marina Caskey<sup>1\*</sup>, Florian Klein<sup>1\*</sup>, Julio C. C. Lorenzi<sup>1</sup>, Michael S. Seaman<sup>2</sup>, Anthony P. West Jr<sup>3</sup>, Noreen Buckley<sup>1</sup>, Gisela Kremer<sup>4,5</sup>, Lilian Nogueira<sup>1</sup>, Malte Braunschweig<sup>1,6</sup>, Johannes F. Scheid<sup>1</sup>, Joshua A. Horwitz<sup>1</sup>, Irina Shimeliovich<sup>1</sup>, Sivan Ben-Avraham<sup>1</sup>, Maggi Witmer-Pack<sup>1</sup>, Martin Platten<sup>4,7</sup>, Clara Lehmann<sup>4,7</sup>, Leah A. Burke<sup>1,8</sup>, Thomas Hawthorne<sup>9</sup>, Robert J. Gorelick<sup>10</sup>, Bruce D. Walker<sup>11</sup>, Tibor Keler<sup>9</sup>, Roy M. Gulick<sup>8</sup>, Gerd Fätkenheuer<sup>4,7</sup>, Sarah J. Schlesinger<sup>1</sup> & Michel C. Nussenzweig<sup>1,12</sup>

**HIV-1 immunotherapy with a combination of first generation monoclonal antibodies was largely ineffective in pre-clinical and clinical settings and was therefore abandoned<sup>1–3</sup>. However, recently developed single-cell-based antibody cloning methods have uncovered a new generation of far more potent broadly neutralizing antibodies to HIV-1 (refs 4, 5). These antibodies can prevent infection and suppress viraemia in humanized mice and nonhuman primates, but their potential for human HIV-1 immunotherapy has not been evaluated<sup>6–10</sup>. Here we report the results of a first-in-man dose escalation phase 1 clinical trial of 3BNC117, a potent human CD4 binding site antibody<sup>11</sup>, in uninfected and HIV-1-infected individuals. 3BNC117 infusion was well tolerated and demonstrated favourable pharmacokinetics. A single 30 mg kg<sup>−1</sup> infusion of 3BNC117 reduced the viral load in HIV-1-infected individuals by 0.8–2.5 log<sub>10</sub> and viraemia remained significantly reduced for 28 days. Emergence of resistant viral strains was variable, with some individuals remaining sensitive to 3BNC117 for a period of 28 days. We conclude that, as a single agent, 3BNC117 is safe and effective in reducing HIV-1 viraemia, and that immunotherapy should be explored as a new modality for HIV-1 prevention, therapy and cure.**

A fraction of HIV-1-infected individuals develop potent neutralizing serologic activity against diverse viral isolates<sup>4,5</sup>. Single-cell cloning methods to isolate antibodies from these individuals<sup>12</sup> revealed that broad and potent neutralization can be achieved by antibodies targeting many sites on the viral envelope<sup>5,13,14</sup>. Many of these antibodies can prevent infection, and some can suppress active infection in humanized mice (hu-mice) or macaques<sup>6–10</sup>. Therefore, it is generally accepted that a vaccine eliciting such antibodies is likely to be protective against HIV-1. However, potent anti-HIV-1 broadly neutralizing antibodies (bNAbs) are highly somatically mutated and many carry other uncommon features such as insertions, deletions, or long complementary determining regions<sup>4,5,11,12,15</sup>, which may account for the difficulty in eliciting such antibodies by immunization. In view of the efficacy of passive bNAb administration in hu-mice and macaques<sup>6–9,16</sup>, it has been suggested that bNAbs should be administered passively, or by viral vectors for prevention and immunotherapy<sup>4,9,16</sup>. However, their safety and efficacy has not been tested in humans.

To determine whether the new generation of more potent bNAbs are safe and active against HIV-1 in humans, we initiated an open label phase 1 study (Fig. 1a) with 3BNC117, an anti-CD4 binding site antibody cloned from a viraemic controller<sup>11</sup>. 3BNC117 neutralizes 195 out of 237 HIV-1 strains comprising 6 different clades with an average

half-maximal inhibitory concentration (IC<sub>50</sub>) of 0.08 µg ml<sup>−1</sup> (Extended Data Fig. 1)<sup>11</sup>. 12 uninfected and 17 HIV-1-infected individuals (Table 1) were administered a single intravenous dose of 1, 3, 10 or 30 mg kg<sup>−1</sup> of 3BNC117 (Extended Data Table 1a). 3BNC117 serum concentrations, plasma HIV-1 viral loads (VL), CD4<sup>+</sup> and CD8<sup>+</sup> T-cell counts, and safety were monitored closely (Fig. 1a, Extended Data Figs 2, 3, and Extended Data Tables 1b, 2). The two groups were comparable for gender, race and age (Table 1).

3BNC117 was generally safe and well tolerated at all doses tested in both uninfected and HIV-1-infected individuals. No grade 3, 4 or serious adverse events and no treatment-related laboratory changes were observed during 56 days of follow up (Extended Data Table 1b). CD4<sup>+</sup> or CD8<sup>+</sup> T-cell counts did not change after 3BNC117 infusion in the HIV-1-infected group, possibly because initial CD4<sup>+</sup> T-cell counts were near normal in most participants (mean absolute CD4<sup>+</sup> T-cell count was 655 cells per µl, Extended Data Fig. 2).

Two different assays were used to measure 3BNC117 levels in serum: TZM.bl neutralization assay to measure activity, and anti-idiotypic-specific ELISA to measure antibody protein levels (Fig. 1b, Extended Data Fig. 3 and Extended Data Tables 4, 5). With few exceptions the two assays were generally in agreement in both groups (Fig. 1b and Extended Data Fig. 3). However, elimination of 3BNC117 activity was more rapid in the HIV-1-infected group, resulting in an estimated average t<sub>1/2</sub> of around 9 days as opposed to around 17 days in uninfected individuals (Fig. 1b and Extended Data Tables 4, 5). We conclude that 3BNC117 has pharmacokinetic properties consistent with a typical human IgG1 in uninfected individuals and a somewhat faster decay rate in HIV-1-viraemic individuals. Similar antigen-dependent enhanced clearance has been reported with anti-cancer antibodies<sup>17</sup>. Although there may be other explanations, we speculate that the increased rate of antibody elimination in the presence of HIV-1 is due to accelerated clearance of antigen–antibody complexes.

Viral loads were measured by standard assays or by single-copy assays. Baseline VLs in HIV-1-infected individuals not on anti-retroviral therapy (ART) varied from 640 to 53,470 copies ml<sup>−1</sup> (mean 9,420 copies ml<sup>−1</sup>). Two participants were on ART at the time of antibody infusion but had detectable baseline VLs (30 and 100 copies ml<sup>−1</sup>). (Table 1, Extended Data Table 2a).

Virologic responses correlated with antibody dose. Infected individuals receiving 1 or 3 mg kg<sup>−1</sup> 3BNC117 doses showed only small and transient changes in viraemia consisting of increases of up to threefold 1 day after infusion, followed by a short temporary decrease, and rapid return to baseline (Fig. 2, Extended Data Fig. 4, and Extended Data

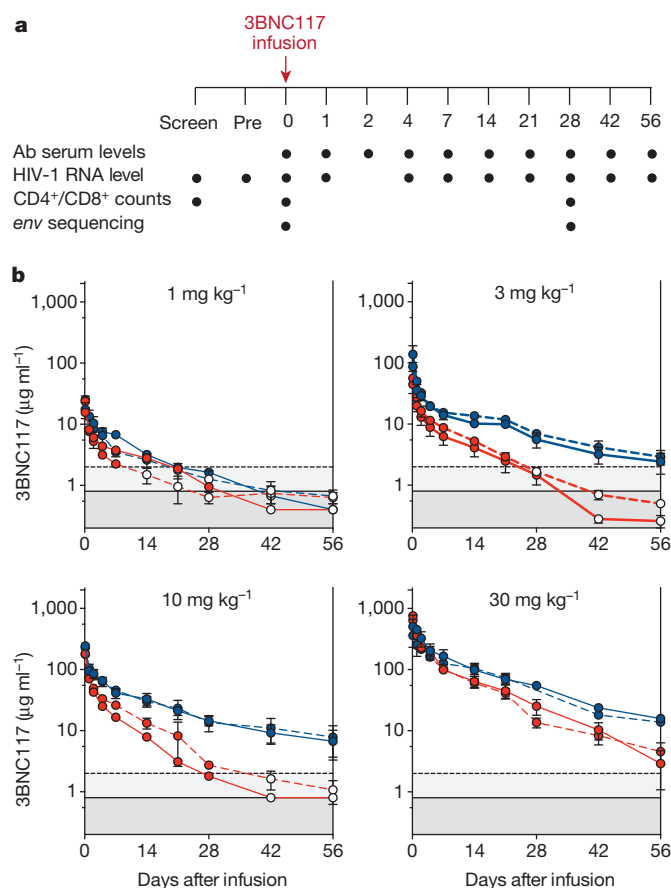
<sup>1</sup>Laboratory of Molecular Immunology, The Rockefeller University, New York, New York 10065, USA. <sup>2</sup>Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA. <sup>3</sup>Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. <sup>4</sup>First Department of Internal Medicine, University Hospital of Cologne, D-50924 Cologne, Germany. <sup>5</sup>Clinical Trials Center Cologne, ZKS Köln, BMBF 01KN1106, University of Cologne, Cologne, Germany. <sup>6</sup>Albert Ludwigs University of Freiburg, 79085 Freiburg, Germany.

<sup>7</sup>German Center for Infection Research (DZIF), partner site Bonn-Cologne, Cologne, Germany. <sup>8</sup>Division of Infectious Diseases, Weill Medical College of Cornell University, New York, New York 10065, USA.

<sup>9</sup>Celldex Therapeutics, Inc., Hampton, New Jersey 08827, USA. <sup>10</sup>AIDS and Cancer Virus Program, Leidos Biomedical Research, Frederick, Frederick National Laboratory for Cancer Research, Frederick, Maryland 21702, USA. <sup>11</sup>Ragon Institute of MGH, MIT and Harvard, Howard Hughes Medical Institute, Massachusetts General Hospital and Harvard Medical School, Cambridge, Massachusetts 02139, USA.

<sup>12</sup>Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10065, USA.

\*These authors contributed equally to this work.



**Figure 1 | Pharmacokinetics of 3BNC117 in uninfected and HIV-1-infected individuals.** **a**, Diagrammatic representation of the study. Time of 3BNC117 infusion indicated by the red arrow, and sampling for 3BNC117 serum levels, HIV-1 viral load, CD4<sup>+</sup>/CD8<sup>+</sup> T cell counts and *env* sequencing as indicated below. **b**, Antibody decay measured in TZM.bl assays (solid lines) and ELISA (dotted lines). Mean values and s.e.m. for uninfected individuals (3 per group) are shown in blue and for HIV-1-infected individuals (2–5 per group) in red. Light grey indicates lower level of accuracy by the ELISA assay and dark grey by the TZM.bl assay. Open circles indicate levels lower than the accuracy threshold.

Table 2a). The magnitude and kinetics of the initial increase in viraemia were consistent with those seen with viral entry inhibitors<sup>18</sup>.

In contrast, 10 out of 11 individuals receiving 10 or 30 mg kg<sup>-1</sup> infusions responded by dropping their VLs by up to 2.5 log<sub>10</sub> (Fig. 2, Extended Data Fig. 4 and Extended Data Table 2a). Two individuals off ART received the 10 mg kg<sup>-1</sup> dose, of whom 1 responded with 1.36

log<sub>10</sub> decline in viraemia and the other did not (Fig. 2 and Extended Data Table 2a). The individual that did not respond was infected with a virus that was completely resistant to 3BNC117 (2C4; IC<sub>50</sub> > 20 μg ml<sup>-1</sup>; Fig. 3 and Extended Data Table 3). All 8 individuals that received the 30 mg kg<sup>-1</sup> dose of 3BNC117 showed highly significant and rapid decreases in their viral loads that varied between -0.8 and -2.5 log<sub>10</sub> (Figs 2 and 3 and Extended Data Table 2a, b). The magnitude of the decline was related to the starting VL and the sensitivity of the subjects' autologous virus to 3BNC117 (Figs 2 and 3, Extended Data Fig. 5). The median time to reach the lowest level in viraemia was 7 days, and the mean drop in VL was 1.48 log<sub>10</sub> at lowest level. When compared to all available pre-treatment measurements, the drop in viraemia was highly significant from days 4 through 28 (Fig. 2 and Extended Data Table 2b). Although the limited data set does not allow us to determine viral set point, 4 of the 8 individuals receiving a single 30 mg kg<sup>-1</sup> infusion did not entirely return to day 0 pre-infusion levels during the observation period of 56 days (Figs 2 and 3, Extended Data Table 2a).

To further examine the virologic effects of 3BNC117 immunotherapy, autologous viral isolates were obtained from cultured PBMCs before (day 0, day -7) and after (day 28) antibody infusion. Paired samples from 12 of the 17 HIV-1-infected individuals were tested for 3BNC117 sensitivity (Fig. 3 and Extended Data Table 3). Samples obtained from individuals infused with 1 mg kg<sup>-1</sup> showed 35- and 13.5-fold decreases in 3BNC117 sensitivity, indicating that the antibody exerts selective pressure on HIV-1 even at the lowest dose (2A3, 2A4; Fig. 3). Similar changes in sensitivity were seen for some (2B1, 2C5) individuals treated with 3 and 10 mg kg<sup>-1</sup>, but others remained 3BNC117-sensitive throughout (2B3) (Fig. 3, Extended Data Table 3). Similarly, at 30 mg kg<sup>-1</sup>, only 2 out of 5 individuals tested showed greater than fivefold reduction in 3BNC117 sensitivity on day 28 (Extended Data Table 3). In contrast, 2C1, 2D1, 2D3 showed only 3.2-, 1.3- and 2.7-fold changes in sensitivity and these individuals did not rebound to baseline viraemia levels at day 28 (Figs 2 and 3). We conclude that, in some individuals, HIV-1 develops high-level resistance to 3BNC117 by 28 days after a single dose, while in others it does not.

To examine the molecular nature of the changes in HIV-1 in response to 3BNC117, we cloned and sequenced HIV-1 envelopes from paired plasma samples from 10 individuals before and 28 days after infusion. Evidence for antibody-induced selection was seen in some but not all samples analysed (Fig. 4, Supplementary Fig. 1). For example, 2C5, who received a 10 mg kg<sup>-1</sup> infusion, selected for a G459D mutation in 15 out of 23 *env* sequences, while the remainder showed a longer V5 loop. The G459D mutation alters the CD4 binding site and can result in resistance to 3BNC117 (ref. 9). Changes in the V5 loop can alter sensitivity to anti-CD4 binding site antibodies by steric clashing with the heavy or light chains of 3BNC117-type antibodies. Similarly, 10 or 30 mg kg<sup>-1</sup> infusions selected single mutations at Q363H (2E1), S461D (2E2), and S274Y (2E2) (Fig. 4, Supplementary Fig. 1). These changes may alter sensitivity to 3BNC117 by interfering with binding<sup>5</sup>. Selection in these 3 individuals is also indicated by the emergence of a distinct group of closely related sequences in phylogenetic trees (Fig. 4, Supplementary Fig. 1). Consistent with the molecular analysis, and the viral culture data (Fig. 3), pseudoviruses produced from serum of 2C5 from days 0 and 28 showed high level 3BNC117 resistance, whereas the changes in pseudoviruses produced from 2C1, 2D1, 2E1 and 2E2 were modest (Fig. 4 and Extended Data Table 3). In contrast, autologous viral isolates from individuals who did not become resistant, or had only small changes in sensitivity, such as 2B3, showed little if any evidence of selection. We conclude that a single infusion of 3BNC117 leads to selection for high-level resistance in some but not all individuals.

Although immunotherapy was initially used to treat infectious diseases, the great majority of therapeutic monoclonal antibodies are currently used to treat cancer and autoimmune diseases. This form of therapy has been shown to be highly effective, well tolerated, and to function in large part by engaging the host immune system through Fc receptors<sup>19</sup>.

In contrast, a role for antibodies in controlling HIV-1 infection has been difficult to establish. For example, the overall course of infection is

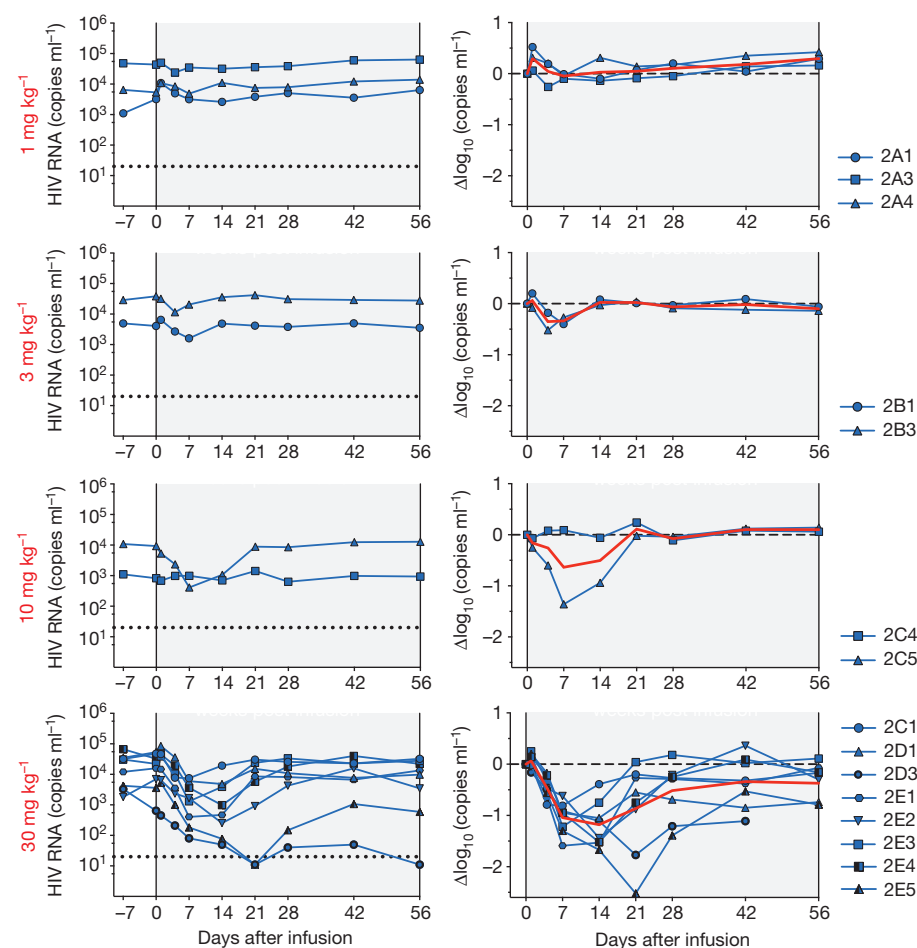
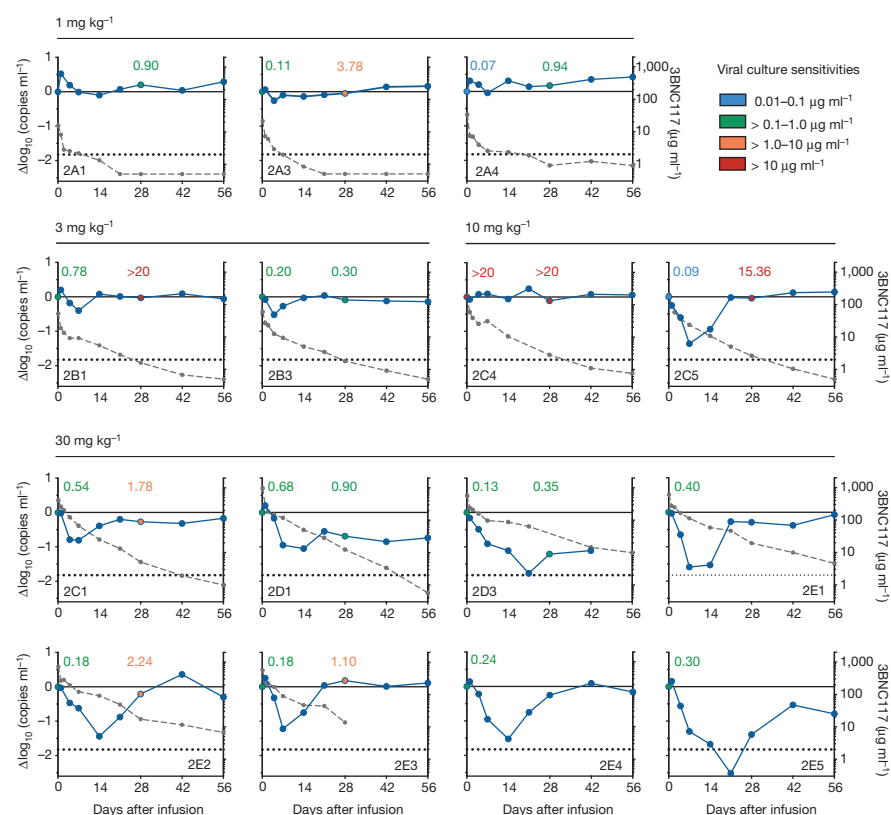
**Table 1 | Study participants demographics**

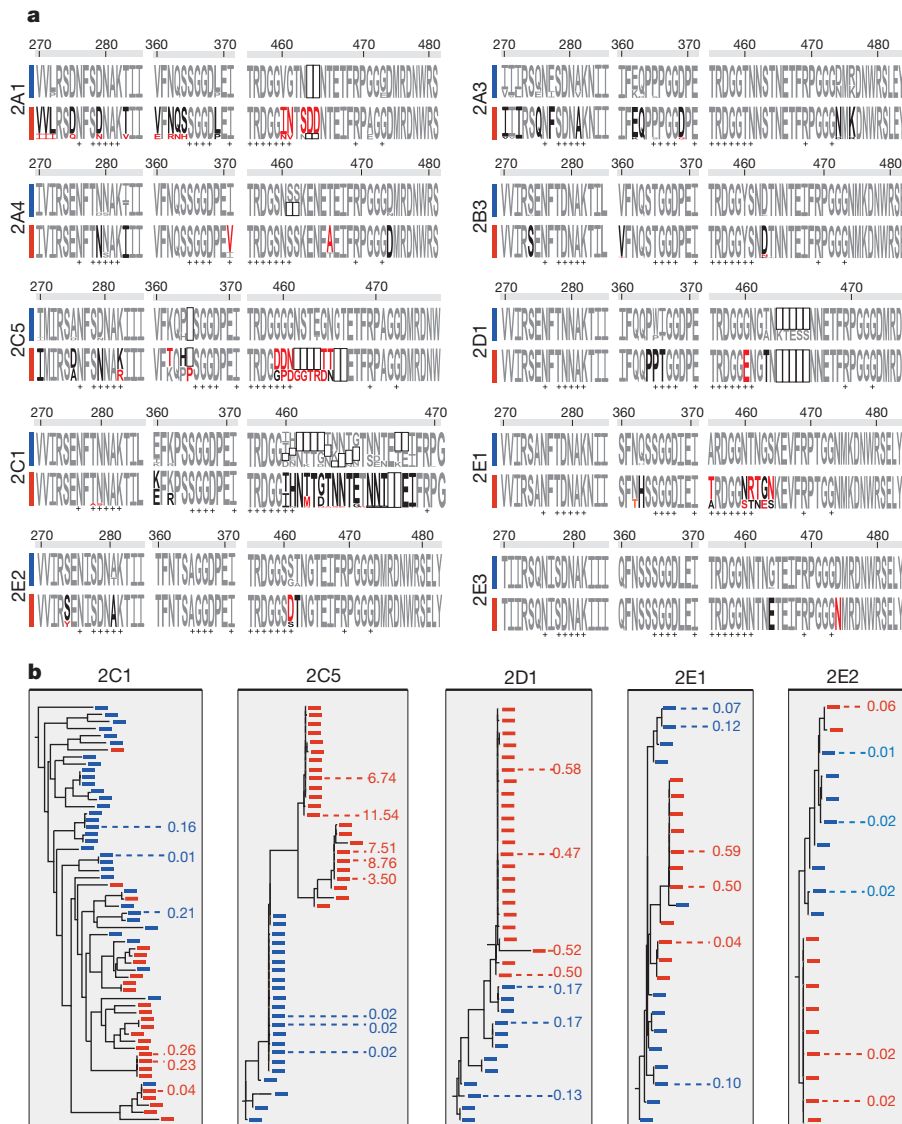
|   | Uninfected (n = 12) | HIV-1-infected (n = 17) |
|---|---------------------|-------------------------|
| Gender (% male)   | 83%                 | 76%                     |
| Mean age (range)  | 43 (22–58)          | 37 (20–54)              |
| Race/ethnicity  |                     |                         |
| White   | 42%                 | 29%                     |
| Black or African American   | 50%                 | 53%                     |
| Hispanic  | 8%                  | 18%                     |
| ART status  |                     |                         |
| On ART n (%)  | —                   | 2 (12%)                 |
| Off ART n (%)   | —                   | 15 (88%)                |
| Mean absolute CD4 <sup>+</sup> count (cells μl <sup>-1</sup> ; day 0) | —                   | 655 (245–1,129)         |
| Mean % CD4 <sup>+</sup> count (day 0)                                 | —                   | 29% (20–42%)            |
| Mean HIV-1 RNA level (copies ml <sup>-1</sup> ; day 0)*               | —                   | 9,420 (640–53,470)      |

\*Mean HIV-1 RNA levels in HIV-1-infected participants off ART.

**Figure 2 | HIV-1 viral load measurements.**

3BNC117 dose indicated in red. Plots (left column) show absolute VLs in HIV-1 RNA copies  $\text{ml}^{-1}$  (y axis) versus time in days after infusion (x axis). Right column shows  $\log_{10}$  changes in VL from day 0. Red line illustrates the average (least-squares means, by mixed-effect linear model). Individual subjects are indicated on the right. Subjects 2E1 to 2E5 were pre-screened for 3BNC117 sensitivity. At  $30 \text{ mg kg}^{-1}$  dose level, the change in viraemia was significant ( $P = 0.004$ ,  $P < 0.001$ ,  $P < 0.001$ ,  $P < 0.011$  at days 4, 7, 14, 21, and 28, respectively) when compared to all available pretreatment values (Extended Data Table 2b).

**Figure 3 | 3BNC117 sensitivity, changes in viraemia, and 3BNC117 levels.** 3BNC117 dose is indicated at the top of the graphs. The left y axis shows  $\log_{10}$  change in viraemia from day 0, and right y axis shows antibody level measured by ELISA. Blue line reflects change in VL and dotted grey line antibody level. Numbers indicate  $\text{IC}_{50}$  values for 3BNC117 of autologous viral isolates measured by TZM.bl assay, colour-coded as indicated in the key. Dotted line indicates lower level of accuracy.



**Figure 4 | HIV-1 envelope sequence analysis after 3BNC117 infusion.**

**a**, HIV-1 envelopes were cloned from plasma samples. Logogram showing *env* gp120 regions (amino acid positions; 270–285, 360–371, and 455 to 471–485, according to HXBc2 numbering) indicating sequence changes from day 0 (blue bar) to day 28 (red bar). The frequency of each amino acid is indicated by its height. Red residues represent mutations that were only found after treatment, black residues represent amino acids that changed in frequency after treatment, empty boxes represent gaps, and + symbols represent 3BNC117 contact sites on gp120 (ref. 30). **b**, Phylogenetic trees show gp120 evolution from day 0 to 28

after treatment for 2C1, 2C5, 2D1, 2E1 and 2E2. Blue and red bars represent sequences obtained on days 0 and 28, respectively (Supplementary Fig. 1). Illustrated values represent the  $IC_{50}$  of 3BNC117 in  $\mu g\ ml^{-1}$  against the cloned HIV-1 pseudoviruses from the analysed sequences (Extended Data Table 3). The geometric means of the pseudoviruses'  $IC_{50}$  values on days 0 and 28, respectively are for 2C1: 0.06 and  $0.14\ \mu g\ ml^{-1}$ ; 2C5: 0.02 and  $7.09\ \mu g\ ml^{-1}$ ; 2D1: 0.15 and  $0.52\ \mu g\ ml^{-1}$ ; 2E1: 0.09 and  $0.23\ \mu g\ ml^{-1}$ ; 2E2: 0.01 and  $0.03\ \mu g\ ml^{-1}$ .

not thought to be altered in individuals that develop bNAbs<sup>20</sup>. Moreover, first generation anti-HIV-1 bNAbs with limited breadth and activity produced little if any measurable effects in hu-mice or viraemic individuals<sup>3,4,21</sup>. However, antibodies can put strong selective pressure on the virus in individuals that develop anti-HIV-1 antibody responses<sup>22–24</sup>. In addition, recent studies in hu-mice showed that, when administered as monotherapy, new generation bNAbs can transiently reduce VLs, and in combination they control viraemia for as long as concentrations remain in the therapeutic range<sup>6,9</sup>. In contrast, single antibodies led to control of viraemia in SHIV-infected macaques for as long as antibody levels remained therapeutic, and immune escape was rarely observed<sup>7,8</sup>. The surprising difference between hu-mice and macaques might be attributed in part to an intact host immune system in the macaque, including endogenous antibodies<sup>25</sup>, or differences between SHIV and HIV-1 infection. Our data establish that passive infusion of single bNAbs can have profound effects on HIV-1 viraemia in humans.

Combinations of antiretroviral drugs are the standard of care for HIV-1 infection because resistance develops to single agents<sup>26</sup>. Similarly, monotherapy with 3BNC117 alone is insufficient to control infection, and we expect that antibody-drug or antibody-antibody combinations will be required for complete viraemic control. Although the current generation of drugs is less expensive than antibodies, the latter have very long half-lives and have the potential to kill infected cells and to enhance host immunity by engaging Fc receptors<sup>19,27</sup>. Moreover, anti-HIV-1 antibodies can be made 100-fold more potent by molecular engineering<sup>28</sup>. Finally, the combination of antibodies with agents that activate latent viruses can interfere with the HIV-1 reservoir in hu-mice<sup>29</sup> and may be critical to HIV-1 eradication strategies.

Given the difficulties in developing an HIV-1 vaccine and in eradicating established infection, passive transfer of monoclonal antibodies is being considered for HIV-1 prevention, therapy, and cure. Our data establish the principle that monoclonal antibodies can be both safe and



effective against HIV-1 in humans. Antibody-mediated immunotherapy differs from currently available drugs in that it has the potential to affect the course of HIV-1 infection by engaging host immunity directly.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 27 January; accepted 19 March 2015.**

**Published online 8 April 2015.**

1. Mehndru, S. *et al.* Adjunctive passive immunotherapy in human immunodeficiency virus type 1-infected individuals treated with antiviral therapy during acute and early infection. *J. Virol.* **81**, 11016–11031 (2007).
2. Trkola, A. *et al.* Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nature Med.* **11**, 615–622 (2005).
3. Armbruster, C. *et al.* Passive immunization with the anti-HIV-1 human monoclonal antibody (hMAb) 4E10 and the hMAb combination 4E10/2F5/2G12. *J. Antimicrob. Chemother.* **54**, 915–920 (2004).
4. Klein, F. *et al.* Antibodies in HIV-1 vaccine development and therapy. *Science* **341**, 1199–1204 (2013).
5. West, A. P. Jr *et al.* Structural insights on the role of antibodies in HIV-1 vaccine and therapy. *Cell* **156**, 633–648 (2014).
6. Klein, F. *et al.* HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* **492**, 118–122 (2012).
7. Barouch, D. H. *et al.* Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. *Nature* **503**, 224–228 (2013).
8. Shingai, M. *et al.* Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature* **503**, 277–280 (2013).
9. Horwitz, J. A. *et al.* HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. *Proc. Natl Acad. Sci. USA* **110**, 16538–16543 (2013).
10. Moldt, B. *et al.* Highly potent HIV-specific antibody neutralization *in vitro* translates into effective protection against mucosal SHIV challenge *in vivo*. *Proc. Natl Acad. Sci. USA* **109**, 18921–18925 (2012).
11. Scheid, J. F. *et al.* Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633–1637 (2011).
12. Scheid, J. F. *et al.* Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458**, 636–640 (2009).
13. Scharf, L. *et al.* Antibody 8ANC195 reveals a site of broad vulnerability on the HIV-1 envelope spike. *Cell Rep.* **7**, 785–795 (2014).
14. Falkowska, E. *et al.* Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. *Immunity* **40**, 657–668 (2014).
15. Klein, F. *et al.* Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* **153**, 126–138 (2013).
16. Balazs, A. B. *et al.* Antibody-based protection against HIV infection by vectored immunoprophylaxis. *Nature* **481**, 81–84 (2012).
17. Glassman, P. M. & Balthasar, J. P. Mechanistic considerations for the use of monoclonal antibodies for cancer therapy. *Cancer Biol. Med.* **11**, 20–33 (2014).
18. Nettles, R. E. *et al.* Pharmacodynamics, safety, and pharmacokinetics of BMS-663068, an oral HIV-1 attachment inhibitor in HIV-1-infected subjects. *J. Infect. Dis.* **206**, 1002–1011 (2012).
19. Nimmerjahn, F. & Ravetch, J. V. Antibody-mediated modulation of immune responses. *Immunol. Rev.* **236**, 265–275 (2010).
20. Euler, Z. *et al.* Cross-reactive neutralizing humoral immunity does not protect from HIV type 1 disease progression. *J. Infect. Dis.* **201**, 1045–1053 (2010).
21. Matsushita, S., Yoshimura, K., Ramirez, K. P., Pisupati, J. & Murakami, T. Passive transfer of neutralizing mAb KD-247 reduces plasma viral load in patients chronically infected with HIV-1. *AIDS* **29**, 453–462 (2015).
22. Liao, H. X. *et al.* Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469–476 (2013).
23. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).
24. Doria-Rose, N. A. *et al.* Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* **509**, 55–62 (2014).
25. Klein, F. *et al.* Enhanced HIV-1 immunotherapy by commonly arising antibodies that target virus escape variants. *J. Exp. Med.* **211**, 2361–2372 (2014).
26. Deeks, S. G., Lewin, S. R. & Havlir, D. V. The end of AIDS: HIV infection as a chronic disease. *Lancet* **382**, 1525–1533 (2013).
27. Bournazos, S. *et al.* Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for *in vivo* activity. *Cell* **158**, 1243–1253 (2014).
28. Diskin, R. *et al.* Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science* **334**, 1289–1293 (2011).
29. Halper-Stromberg, A. *et al.* Broadly neutralizing antibodies and viral inducers decrease rebound from HIV-1 latent reservoirs in humanized mice. *Cell* **158**, 989–999 (2014).
30. Zhou, T. *et al.* Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* **39**, 245–258 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank all individuals who participated in this study. We thank the Rockefeller University Hospital Clinical Research Support Office and nursing staff, as well as T. Kümmerle, C. Wyen, and L. Siegel for help with recruitment; S. Kiss for ophthalmologic assessments; F. Maldarelli and J. Lifson for single-copy analysis and B. Freemire for technical assistance; J. Pring, A. Almakari, C. Unsen, S. Hadrigan, E. Thomas, H. Gruell, D. Gillor, and U. Sandaradura de Silva for sample processing and study coordination, and N. Rodziewicz for help with HIV-1 culturing. We thank A. Louie and C. Conrad for help with regulatory submissions; L. Thomas for IND-enabling studies, L. Vitale for cell line and anti-idiotypic antibody development, B. Riordan, A. Rayo and J. Andreozzi for anti-idiotypic ELISA method development and sample analysis, R. Hammond for process development, and S. DiSciullo for 3BNC117 manufacturing; J. Pery for performing neutralization assays. We thank M. Suarez-Farinas for support with statistical analysis; P. Fast and H. Park for clinical monitoring; and E. Gotschlich and B. Collier for input on study design. J.C.C.L. is supported by an award from CNPq “Ciencia sem Fronteiras” Brazil (248676/2013-0). This work was supported in part by the Bill and Melinda Gates Foundation Collaboration for AIDS Vaccine Discovery (CAVD) Grants OPP1033115 (M.C.N.), OPP1092074 (M.C.N.), OPP1040753 (A.P.W.) and OPP1032144 (M.S.S.), and U19AI111825-01 Cooperative Centers on Human Immunology from NIH to M.C.N., by grant #U11 TR000043 from the National Center for Advancing Translational Sciences (NCATS), by a grant from the Robertson Foundation to M.C.N., in part with Federal funds from the NCI/NIH, under Contract No. HHSN261200800001E, and a grant from the German Center for Infection Research (DZIF) to G.F. 3BNC117 was generated from a subject in the International HIV Controller Study, supported by the Mark and Lisa Schwartz Foundation and CAVD Grant 43307. M.C.N. and B.D.W. are Howard Hughes Medical Institute Investigators. The authors declare no competing financial interests.

**Author Contributions** M.C. and F.K. planned and implemented the study, analysed the data, and wrote the manuscript; J.C.C.L. performed sequence analyses, and contributed to writing the manuscript; M.S.S. performed TZM-bl neutralization assays; A.P.W. assisted with sequence analyses; N.B., G.K., S.B.-A., M.W.-P., M.P., L.A.B. implemented the study; L.N. and M.B. performed cloning and sequencing; I.S., C.L. coordinated sample processing; T.H. performed ELISA assays; and R.J.G. performed single copy assays. T.K. was responsible for 3BNC117 manufacture and provided regulatory guidance; J.F.S., B.D.W., J.A.H. contributed to study design and helped with the manuscript; R.M.G. contributed to study design; and G.F. and S.J.S. contributed to study design and implementation. M.C.N. planned and implemented the study, analysed the data and wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.C.N. ([nussen@rockefeller.edu](mailto:nussen@rockefeller.edu)).

## METHODS

**Study design.** An open-label, dose-escalation phase 1 study was conducted in uninfected (Group 1) and HIV-1-infected subjects (Group 2; <http://www.clinicaltrials.gov>; NCT02018510). Study participants were enrolled sequentially according to eligibility criteria. A standard “3+3” phase I trial design was used in the dose-escalation phase of the study. 3BNC117 was administered as a single intravenous infusion at four dose levels: 1 mg kg<sup>-1</sup> (subjects 1A1, 1A2, 1A3, 2A1, 2A3, 2A4), 3 mg kg<sup>-1</sup> (subjects 1B1, 1B2, 1B3, 2B1, 2B2, 2B3), 10 mg kg<sup>-1</sup> (subjects 1C1, 1C3, 1C4, 2C2, 2C4, 2C5) or 30 mg kg<sup>-1</sup> (subjects 1E2, 1E3, 1E5, 2D1, 2C1, 2D3, 2E1, 2E2, 2E3, 2E4, 2E5), at a rate of 100 or 250 ml h<sup>-1</sup>. All participants provided written informed consent before participation in the study and the study was conducted in accordance with Good Clinical Practice. The protocol was approved by the Federal Drug Administration in the USA, the Paul Ehrlich Institute in Germany and the Institutional Review Boards at the Rockefeller University and the University of Cologne.

**Study participants.** All study participants were recruited at the Rockefeller University Hospital, New York, USA and at the University Hospital Cologne, Cologne, Germany. Eligible subjects were adults aged 18–65 years, HIV-1-infected or uninfected, and without concomitant hepatitis B or C infections. HIV-1-infected subjects enrolled in study groups 2A through 2E were ART-experienced or naive. In groups 2A through 2D, subjects were either off standard ART for at least 8 weeks before study participation and had plasma HIV-1 RNA levels between 2,000 and 100,000 copies ml<sup>-1</sup>, or they were on standard ART but had plasma HIV-1 RNA levels >20 copies ml<sup>-1</sup>, measured on two separate occasions at least 1 week apart. Subjects recruited into group 2E were HIV-1-infected, off ART (2,000–100,000 copies ml<sup>-1</sup>), and differed from the other groups in that they were pre-screened for sensitivity to 3BNC117 as described below. Subjects with CD4<sup>+</sup> T-cell counts <300 cells µl<sup>-1</sup>, clinically relevant deviations from normal physical findings, abnormal electrocardiogram (ECG), and/or laboratory examinations were excluded. Women of childbearing potential were required to have a negative result of a serum pregnancy test on the day of 3BNC117 infusion. HIV-1-infected individuals, who were not on standard ART at enrolment, were given the option to initiate ART 6 weeks after 3BNC117 infusion.

**Study procedures.** The appropriate volume of 3BNC117 was calculated according to study dose group, diluted in sterile normal saline to a total volume of 100 or 250 ml, and administered intravenously over 60 min. Study participants received 3BNC117 on day 0 and remained under close monitoring in the inpatient unit of the Rockefeller University Hospital for 24 h. Participants returned for frequent follow up visits for safety assessments that included physical examination, measurement of clinical laboratory parameters such as haematology, chemistry, urinalysis, coagulation times, pregnancy tests (for women) as well as HIV-1 viral loads and CD4<sup>+</sup> and CD8<sup>+</sup> T-cell counts (Fig. 1a). Adverse events were graded according to the DAIDS AE Grading Table (HIV-1-infected groups) or the Toxicity Grading Scale for Healthy Adult and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials (uninfected groups). Blood samples (30 to 120 ml) were collected before and at multiple times after 3BNC117 infusion. Samples were processed within 4 h of collection, and serum and plasma samples were stored at -80 °C. PBMCs were isolated by density gradient centrifugation. The absolute number of peripheral blood mononuclear cells was determined by an automated cell counter (Vi-Cell XR; Beckman Coulter), and cells were cryopreserved in fetal bovine serum plus 10% DMSO.

**Plasma HIV-1 RNA levels.** Plasma was collected for measuring HIV-1 RNA levels at screening (from day -49 to day -14), the day -7 pre infusion visit (from day -42 to day -2), day 0 (before infusion), and on days 1, 4, 7, 14, 21, 28, 42 and 56. HIV-1 RNA levels were determined using the Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 Assay, Version 2.0, which detects 20 to 10 × 10<sup>6</sup> copies ml<sup>-1</sup>, or by the ABBOTT RealTime Assay, which detects 40 to 10 × 10<sup>6</sup> copies ml<sup>-1</sup>. In samples with HIV-1 RNA <20 copies ml<sup>-1</sup>, viraemia was measured by a quantitative real-time, reverse transcriptase (RT)-initiated PCR (RT-PCR) assay that can quantify HIV-1 RNA down to 1 copy ml<sup>-1</sup> as previously described<sup>31</sup>.

**CD4<sup>+</sup> and CD8<sup>+</sup> T cells.** CD4<sup>+</sup> and CD8<sup>+</sup> T-cell counts were determined at screening, on day 0 (before infusion), and day 28 by a clinical flow cytometry assay, performed at LabCorp or at the University Hospital Cologne. Leukocytes were determined as CD45<sup>+</sup> cells. Percentage of cells positively stained for CD3, CD4, CD8 as well as the CD4/CD8 ratio were analysed with the BD Multiset software (BD Biosciences).

**3BNC117 study drug.** 3BNC117 is a recombinant, fully human IgG1κ mAb recognizing the CD4 binding site on the HIV-1 envelope<sup>11</sup>. The antibody was cloned from an HIV-1-infected viraemic controller in the International HIV Controller Study<sup>11,32</sup>, expressed in Chinese hamster ovary cells (clone 5D5-5C10), and purified using standard methods. The 3BNC117 drug substance was produced at Celldex Therapeutics Fall River (MA) GMP facility, and the drug product was fill-finished at Gallus BioPharmaceuticals (NJ). The resulting purified 3BNC117 was supplied as a single use sterile 20 mg ml<sup>-1</sup> solution for intravenous

injection in 8.06 mM sodium phosphate, 1.47 mM potassium phosphate, 136.9 mM sodium chloride, 2.68 mM potassium chloride, and 0.01% polysorbate 80. 3BNC117 vials were shipped and stored at 4 °C.

**Measurement of 3BNC117 serum levels.** Serum levels of 3BNC117 were determined by using two separate methods (ELISA and TZM.bl). 3BNC117 serum concentrations were measured by a validated sandwich ELISA. Plates (Sigma-Aldrich PN: CLS3590 96-well, High Bind, polystyrene) were coated with 4 µg ml<sup>-1</sup> of an anti-idiotypic antibody specifically recognizing 3BNC117 (anti-ID 1F1 mAb), and incubated overnight at 2–8 °C. After washing, plates were blocked for 1 h with 5% BSA. Serum samples, QCs and standards were added (1:50 minimum dilution in 5% BSA) and incubated for 1 h at room temperature. 3BNC117 was detected using an horseradish peroxidase (HRP)-conjugated mouse anti-human IgG kappa-chain-specific antibody (Abcam PN: ab79115) and the HRP substrate tetra-methylbenzidine. 3BNC117 concentrations were then interpolated from a standard curve of 3BNC117 using a 4 parameter logistic curve-fitting algorithm. The reference standard and positive controls were created from the drug product lot of 3BNC117 used in the clinical study. The capture anti-idiotypic antibody was produced by immunizing BALB/c mice with a Fab' fragment of 3BNC117 and plasma was tested for the presence of neutralizing antibodies in an ELISA. Briefly, the HIV-1 antigen (2CC core protein) was coated to a plate and blocked. Plasma dilutions were pre-incubated with a sub-saturating concentration of 3BNC117 then added to the plate. Binding to the antigen was detected with an HRP-conjugated goat anti-mouse IgG-Fc specific antibody. Plasma was considered neutralizing if it was able to block the binding of 3BNC117 to the antigen coated plate. Two mice were selected for fusion. Those hybridomas that showed high specificity when comparing binding to 3BNC117 versus binding to the irrelevant human IgG1 antibody were selected to expand, screened in the neutralization assay as described above, subcloned and purified for use in the anti-idiotypic-specific ELISA.

In addition, the concentration of active 3BNC117 was determined by TZM.bl neutralization assay<sup>7,8</sup>. Serum samples were heat-inactivated for 1 h at 56 °C and measured for neutralizing activity against an HIV-1 strain that was highly sensitive to 3BNC117 but resistant to any autologous HIV-1 neutralizing serum activity. In all uninfected subjects serum samples were tested against Q769.d22 and ID<sub>50</sub> values were derived by using a 5-parameter curve fitting, considering accurate within the pre-established limits (threefold variation with a 20% error rate). The serum concentration of active 3BNC117 was calculated by taking into account the sera ID<sub>50</sub> titres multiplied by the known IC<sub>50</sub> of 3BNC117 for Q769.d22. In HIV-1-infected subjects pre-infusion samples were first tested against a panel of 3BNC117-sensitive HIV-1 strains that included Q769.d22 (Clade A1, Tier 2), YU2.DG (Clade B, Tier 2), Q259.d2.17 (Clade C, Tier 1B), Q842.d12 (Clade A1, Tier 2), ZM135M.PL10a (Clade C, Tier 2), and TRO.11 (Clade B, Tier 2). A single strain per subject was selected that showed no or only minimal background activity and 3BNC117 serum levels were determined in the same way as described for Q769.d22.

**Pharmacokinetic analysis.** Blood samples were collected immediately before, at the end, 0.5, 3, 6, 9, 12 and 24 h after completion of the 3BNC117 infusion, and on days 2, 4, 7, 14, 21, 28, 42 and 56. 3BNC117 serum levels were obtained from ELISA (Celldex Therapeutics) and TZM.bl neutralization assay, and PK-parameters were estimated by performing a non-compartmental analysis (NCA) using WinNonlin 6.3.

**Neutralization assay.** Serum samples, viral supernatants, and control antibodies were tested against HIV-1 envelope pseudoviruses as previously described<sup>33,34</sup>.

**Virus cultures.** Autologous virus was retrieved from HIV-1 infected individuals as previously described<sup>35</sup>. Briefly, healthy donor peripheral blood mononuclear cells (PBMCs) were obtained by leukapheresis from a single donor. Cells were cultured at a concentration of 5 × 10<sup>6</sup> ml<sup>-1</sup> in Iscove's modified Dulbecco's medium (IMDM; Gibco) supplemented with 10% fetal bovine serum (FBS; HyClone, Thermo Scientific), 1% penicillin/streptomycin (Gibco), and 1 µg ml<sup>-1</sup> phytohaemagglutinin (Life Technologies) at 37 °C and 5% CO<sub>2</sub>. After 2–3 days, 5 × 10<sup>6</sup> cells were transferred into IMDM supplemented with 10% FBS, 1% penicillin/streptomycin, 5 µg ml<sup>-1</sup> polybrene (Sigma), and 10 U ml<sup>-1</sup> of IL-2, and co-incubated with 2–3 × 10<sup>6</sup> PBMCs from the study participants obtained before and 28 days or later after 3BNC117 infusion (Extended Data Table 3). Media was replaced on a weekly basis and culture supernatants quantified using the Alliance HIV-1 p24 Antigen ELISA kit (PerkinElmer) according to the manufacturer's instructions. Tissue culture infectious dose for 50% of cells (TCID<sub>50</sub>) values were determined for all HIV-1 containing supernatants<sup>33,34</sup> and then tested for sensitivity against 3BNC117 in a TZM.bl neutralization assay. Outgrowth of autologous virus isolates from subjects with low level viraemia was performed as previously described<sup>36</sup>. Briefly, 2–5 million CD4<sup>+</sup> T cells were cultured in the presence of 10 million irradiated healthy donor PBMCs and 3 million healthy donor PHA stimulated CD8<sup>+</sup> depleted lymphoblasts. Lymphoblasts were replenished weekly by adding 3

million healthy donor PHA stimulated CD8<sup>+</sup> depleted lymphoblasts. Blood samples and leukopheresis were collected under separate IRB-approved protocols and after study participants provided informed consent.

**Sequence analysis.** HIV-1 RNA was extracted from plasma samples using the Qiagen MinElute Virus Spin kit (Qiagen) followed by first strand cDNA synthesis using SuperScript III reverse transcriptase (Invitrogen Life Technologies) and the antisense primer env3out 5'-TTGCTACTTGTGATTGCTCCATGT-3'<sup>37</sup> or 5'-GGTGTGTAGTTCTGCCAATCAGGGAAGWAGCCTTGTG-3'<sup>6</sup>. gp160 *env* was amplified using envB5out 5'-TAGAGCCCTGGAAGCATCCAGGAAG-3' and envB3out 5'-TTGCTACTTGTGATTGCTCCATGT-3' in the first round and second round nested primers envB5in 5'-CACCTTAGGCATCTCCTATG-GCAGGAAGAAG-3' and envB3in 5'-GTCTCGAGATACTGCTCCACCC-3'. First round PCR was performed using a High Fidelity Platinum Taq (Invitrogen) at 94 °C, 2 min; (94 °C, 15 s; 55 °C 30 s; 68 °C, 4 min) × 35; 68 °C, 15 min. Second round PCR was performed with 2 µl of 1. PCR product as template and Phusion Hot Start Polymerase at 98 °C, 30 s; (98 °C, 8 s; 55 °C, 20 s; 72 °C, 1 min) × 35; 72 °C, 6 min. gp120 *env* was amplified using first round primers and conditions as described, but second round was performed by using second round nested primers 5'-TAGAAAGAGCAGAAGACAGTGGCAATGA-3' and 5'-TCATCAATGGTGGTGTGATGATGATGTTTTCTCTCTGCACCACTCTTCT-3'. Second round PCR was performed with 1 µl of 1. PCR product as template and High Fidelity Platinum Taq (Invitrogen) at cycling conditions, 94 °C, 2 min; (94 °C, 15 s; 58 °C, 35 s; 68 °C, 2 min and 30 s) × 35; 68 °C, 10 min. Following the second-round PCR amplification, 0.5 µl Taq polymerase was added to each 50 µl reaction and an additional 72 °C extension for 15 min was performed to add 3' dA overhangs for cloning inserts into pCR4-TOPO. PCR amplicons were gel-purified and ligated into pCR4-TOPO (Invitrogen) or pcDNA3.1 (Invitrogen), followed by transfection into MAX Efficiency Stbl2 Competent Cells (Life Technologies). Individual colonies were analysed for insert length by PCR, and successfully cloned envelopes sequenced by a set of *env* specific primers. Sequence alignments and mutation analysis of gp120 and gp160 was performed by using Geneious Pro software, version 5.6.7 (Biomatters Ltd), and residues were numbered according to HXBc2. Phylogenetic trees were generated using the PhyML tool at the Los Alamos HIV website<sup>38</sup>, and sequence analysis was performed using Antibody database by Anthony West<sup>39</sup>. Logograms were generated using the Weblogo 3.0 tool<sup>40</sup>. Selected sequences were used to generate pseudoviruses and tested for 3BNC117 sensitivity in a TZM.bl assay<sup>37</sup>.

**Statistical analyses.** The sample size to detect >1 log<sub>10</sub> decline in viraemia with 80% power at 5% of significance was determined to be 5 HIV-1-infected individuals, not on ART, infected with 3BNC117-sensitive viruses assuming that the standard deviation would be similar to 3BNC117 effects in hu-mice<sup>9</sup>. Adverse events were summarized by the number of subjects who experienced the event,

by severity grade and by relationship to 3BNC117 according to the DAIDS AE Grading Table (HIV-1-infected groups) or the Toxicity Grading Scale for Healthy Adult and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials (uninfected groups). PK-parameters were estimated by performing a non-compartmental analysis (NCA) using WinNonlin 6.3. CD4<sup>+</sup> and CD8<sup>+</sup> T-cell counts before and after 3BNC117 were analysed by one-way ANOVA. To assess the changes in HIV-1 viral loads, we used a mixed-effect linear model where dose and time were fixed effects and random intercepts for each participant. These models take full advantage of the repeated measure structure of the data while estimating parameters for each dose simultaneously, hence improving the power of small study groups. The final model was fitted assuming an AR(1) correlation structure over time, which was the best in terms of AIC/BIC criteria. The significance of the effect of 3BNC117 on viral load, defined as change between each time point and day 0, was assessed using least-squares means within each dose group (Extended Data Table 2b). Sensitivity analysis was also carried out with variations of these models and the same conclusions were achieved. Pearson and nonparametric Spearman coefficients were calculated to assess the correlation between maximum drop in viraemia after 3BNC117 infusion and baseline HIV-1 viral load or baseline sensitivity of autologous viruses to 3BNC117.

31. Somsouk, M. *et al.* The immunologic effects of mesalamine in treated HIV-infected individuals with incomplete CD4<sup>+</sup> T cell recovery: a randomized crossover trial. *PLoS ONE* **9**, e116306 (2014).
32. Pereyra, F. *et al.* The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science* **330**, 1551–1557 (2010).
33. Montefiori, D. C. Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. *Curr. Protoc. Immunol.* **12**, Unit 12.11 (2005).
34. Li, M. *et al.* Human immunodeficiency virus type 1 *env* clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J. Virol.* **79**, 10108–10125 (2005).
35. van 't Wout, A. B., Schuitemaker, H. & Kootstra, N. A. Isolation and propagation of HIV-1 on peripheral blood mononuclear cells. *Nature Protocols* **3**, 363–370 (2008).
36. Laird, G. M. *et al.* Rapid quantification of the latent reservoir for HIV-1 using a viral outgrowth assay. *PLoS Pathog.* **9**, e1003398 (2013).
37. Salazar-Gonzalez, J. F. *et al.* Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J. Virol.* **82**, 3952–3970 (2008).
38. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
39. West, A. P. Jr. Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. *Proc. Natl Acad. Sci. USA* **110**, 10598–10603 (2013).
40. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

# a

## 3BNC117

| Clade             | number of tested strains | Mean IC <sub>50</sub> <sup>*</sup> | Coverage (%)<br>(IC <sub>50</sub> <10 µg/ml) |
|-------------------|--------------------------|------------------------------------|--|
| A                 | 57                       | 0.04                               | 86.0   |
| AE                | 22                       | 0.08                               | 90.9   |
| B                 | 49                       | 0.08                               | 91.8   |
| C                 | 84                       | 0.14                               | 71.4   |
| D                 | 17                       | 0.11                               | 88.2   |
| G                 | 7                        | 0.16                               | 71.4   |
| <b>All clades</b> | <b>237**</b>             | <b>0.08</b>                        | <b>82.3</b>                                  |

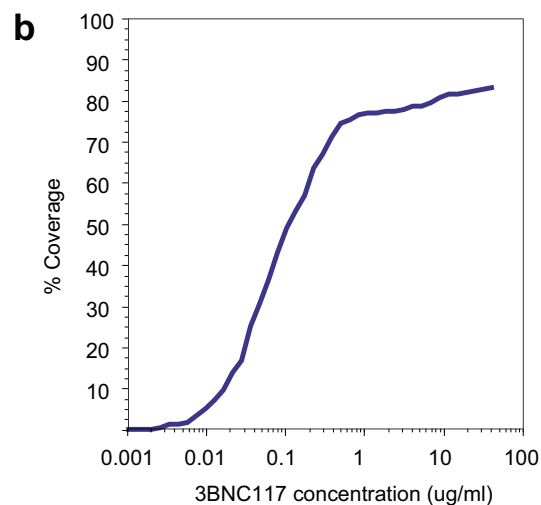
Data retrieved from AntibodyDatabase (West *et al.*, PNAS, 2013)

\* Geometric Mean

\*\* Includes 1 strain, of which no clade was determined

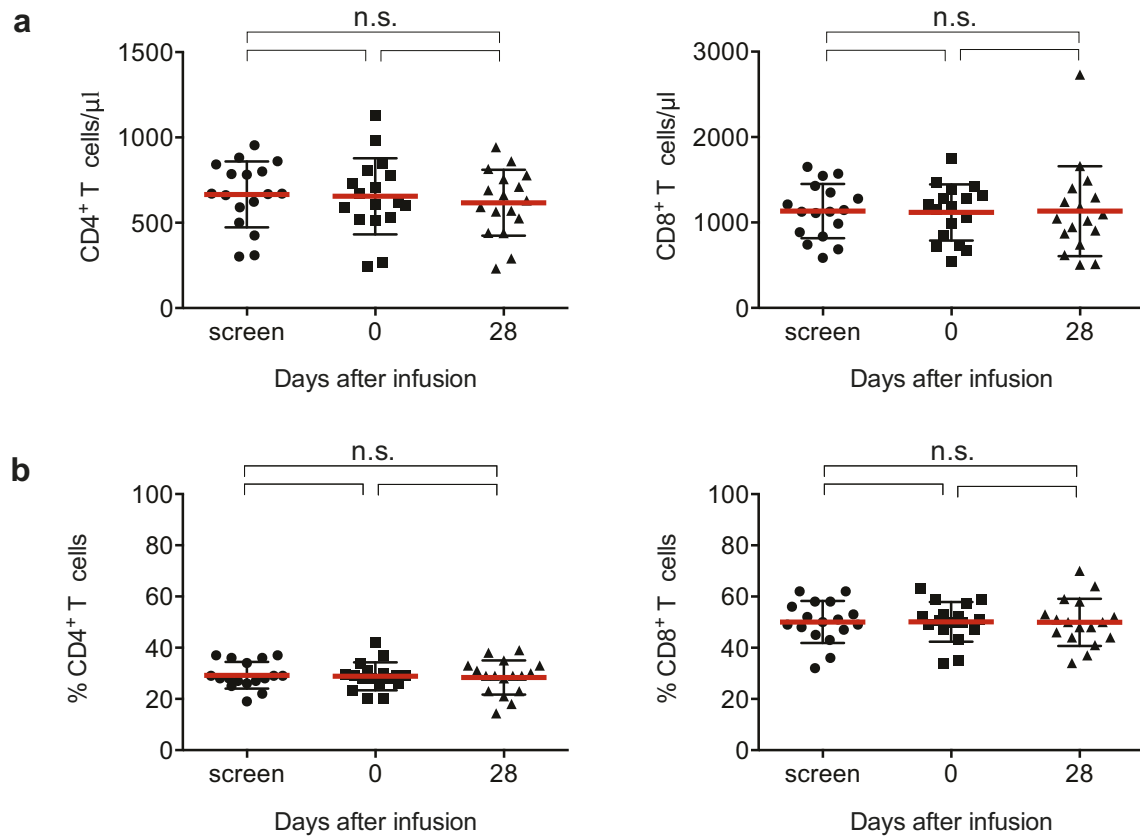
## Extended Data Figure 1 | HIV-1 neutralizing activity of 3BNC117.

**a**, Summary of 3BNC117 neutralizing *in vitro* activity based on 237 HIV-1 isolates comprising 6 different clades. Data were retrieved from the



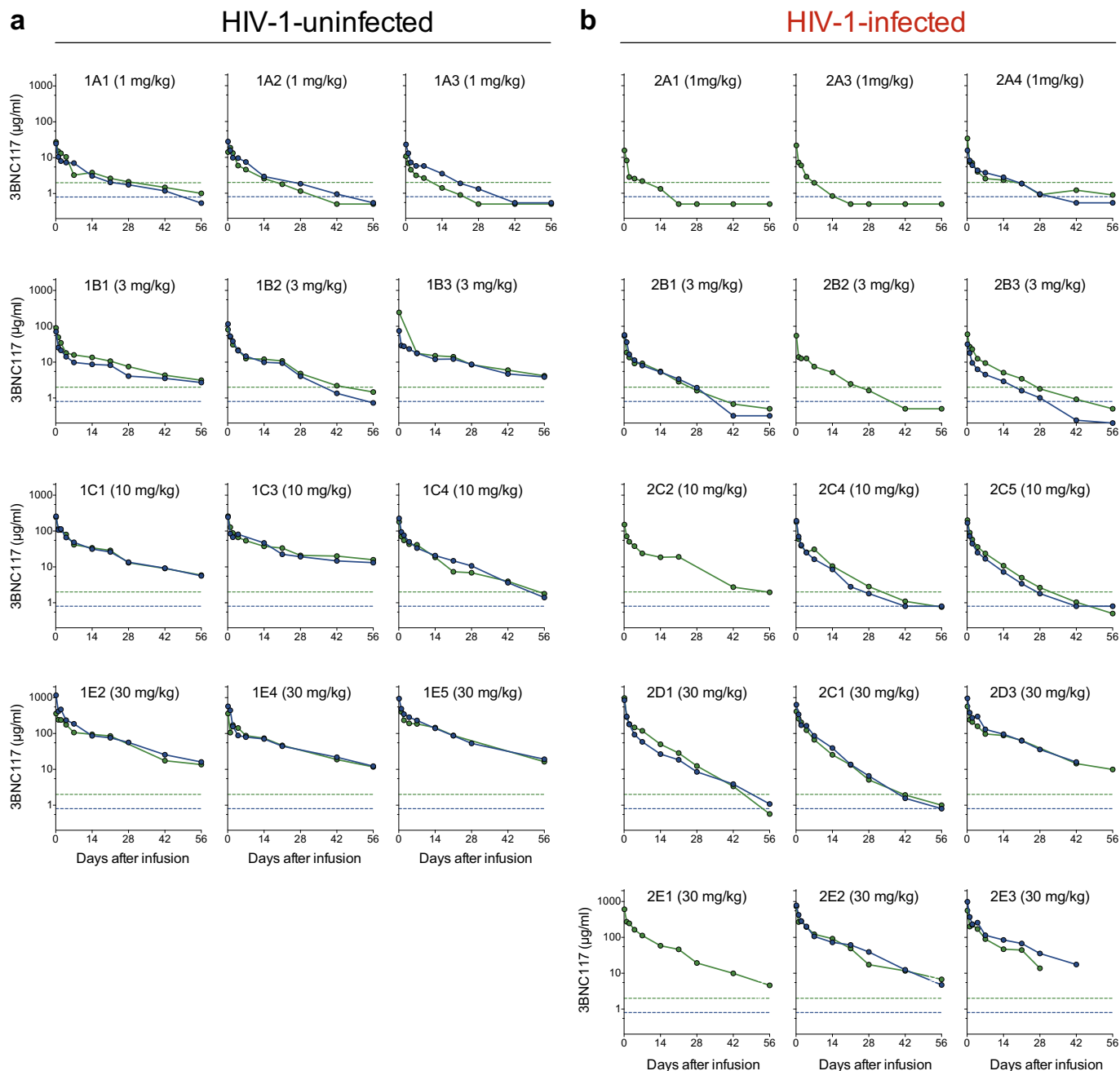
‘AntibodyDatabase’ by A.P.W. (ref. 39). **b**, Illustration of the fraction (that is, % coverage; *y* axis) of HIV-1 isolates that are neutralized at a given IC<sub>50</sub> (µg ml<sup>-1</sup>; *x* axis) using the same data set.





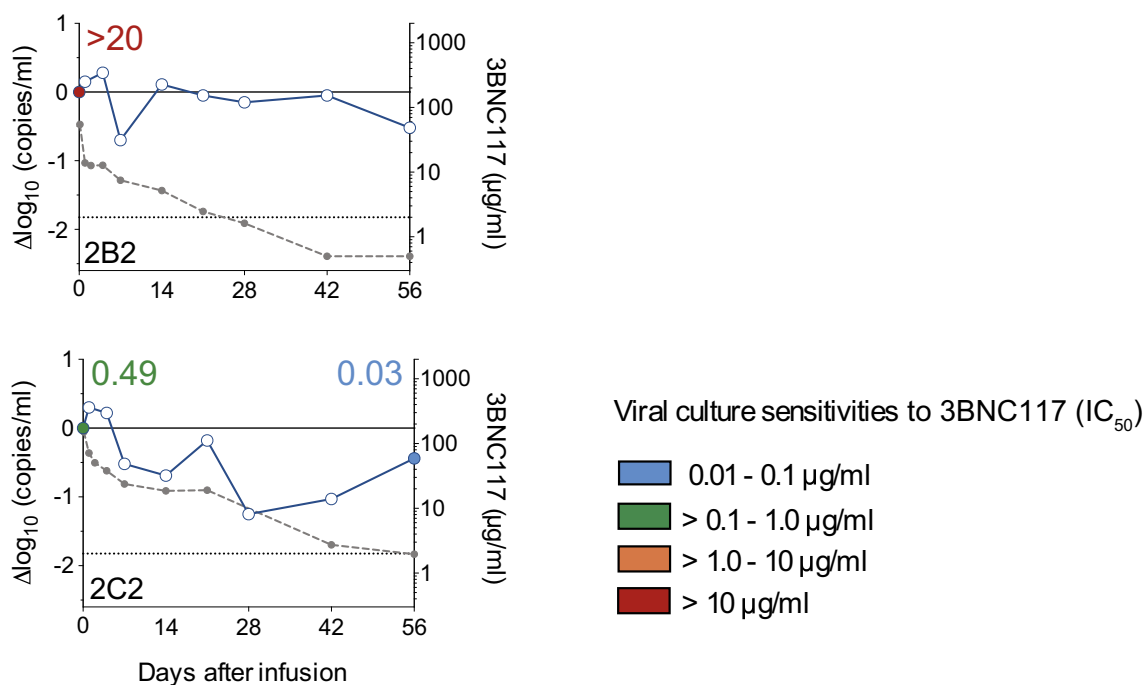
**Extended Data Figure 2 | CD4<sup>+</sup> and CD8<sup>+</sup> T-cell counts before and after 3BNC117 infusion.** **a**, Absolute numbers (cells  $\mu$ l<sup>-1</sup>) of CD4<sup>+</sup> and CD8<sup>+</sup> T-cell counts of all enrolled HIV-1-infected participants at screen, on the day of 3BNC117 infusion (day 0), and at day 28 after infusion. **b**, Percentage of CD4<sup>+</sup>

and CD8<sup>+</sup> T cells for the same subjects and time points. Mean and standard deviation are indicated in red and black, respectively. No significant differences between pre-infusion and post infusion levels were detected by using one-way ANOVA.



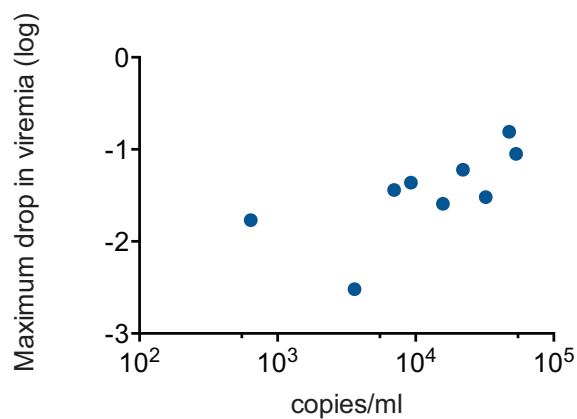
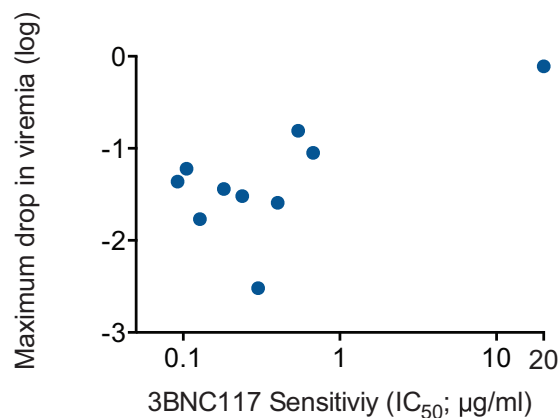
**Extended Data Figure 3 | 3BNC117 serum concentration and activity in single subjects. a, b,** Serum levels of 3BNC117 in all uninfected (a) and HIV-1-infected (b) individuals that received 1, 3, 10, or 30 mg kg<sup>-1</sup> 3BNC117 at day 0.

Antibody levels were measured by a sandwich ELISA using an anti-3BNC117 specific antibody (green) or by measuring the 3BNC117 serum activity in a T2M.b1 neutralization assay (blue).



**Extended Data Figure 4 | 3BNC117 sensitivity and changes in viraemia in 2 ART-treated subjects.** Both subjects were on ART when enrolled in the study and received a single dose (2B2, 3  $\text{mg kg}^{-1}$ ; 2C2, 10  $\text{mg kg}^{-1}$ ) of 3BNC117 at day 0. The left y axis shows  $\log_{10}$  change in viraemia from baseline, and right y

axis shows antibody level measured in ELISA. Blue line reflects change in VL and dotted grey line antibody level. Numbers indicate  $IC_{50}$  values for 3BNC117 of autologous viral isolates measured by TZM.bl assay, colour-coded as indicated on the right.

**a** Starting viremia versus maximum drop in viral load**b** Drop in viremia versus 3BNC117 sensitivity

**Extended Data Figure 5 | Correlating viral decay with 3BNC117 sensitivity and starting viral load.** **a**, Maximum decline in viral load in ART-untreated HIV-1-infected participants with baseline 3BNC117-sensitive viruses ( $IC_{50} < 1 \mu g \text{ ml}^{-1}$ ) versus pre-treatment (day 0) viral load (Pearson coefficient  $r = 0.72$   $P = 0.03$ ; Spearman coefficient  $\rho = 0.78$ ,  $P = 0.02$ ). **b**, Maximum

drop in viral load in HIV-1-infected and viremic individuals receiving a 10 or 30  $\text{mg kg}^{-1}$  dose of 3BNC117 ( $y$  axis) versus baseline autologous virus sensitivity to 3BNC117 ( $x$  axis; Pearson coefficient  $r = 0.69$   $P = 0.03$ ; Spearman coefficient  $\rho = 0.41$ ,  $P = 0.23$ ).



Extended Data Table 1 | Baseline characteristics of HIV-1-infected individuals and 3BNC117 safety data

**a**

| Study ID | 3BNC117 dose | Age | Years since HIV Diagnosis | Current ART regimen | Clade | HIV-RNA level (copies/ml) | abs. CD4 <sup>+</sup> T cell count (day 0; cells/mm <sup>3</sup> ) |
|----------|--------------|-----|---------------------------|---------------------|-------|---------------------------|--|
| 2A1      | 1 mg/kg      | 35  | 11                        | ART naïve           | B     | 3,210                     | 674  |
| 2A3      | 1 mg/kg      | 39  | 14                        | Off ART             | B     | 43,650                    | 520  |
| 2A4      | 1 mg/kg      | 42  | 8                         | ART naïve           | B     | 5,340                     | 607  |
| 2B1      | 3 mg/kg      | 20  | 1                         | Off ART             | ND    | 4,090                     | 264*   |
| 2B2      | 3 mg/kg      | 48  | 20                        | DRV/r/TDF/FTC**     | ND    | 100                       | 706  |
| 2B3      | 3 mg/kg      | 20  | 1                         | ART naïve           | B     | 38,190                    | 777  |
| 2C2      | 10 mg/kg     | 51  | 12                        | ATV/r/3TC/ZDV**     | ND    | 30                        | 728  |
| 2C4      | 10 mg/kg     | 54  | 23                        | Off ART             | ND    | 820                       | 805  |
| 2C5      | 10 mg/kg     | 50  | 4                         | ART naïve           | B     | 9,260                     | 245*   |
| 2D1      | 30 mg/kg     | 33  | 3                         | ART naïve           | B     | 53,470                    | 980  |
| 2C1      | 30 mg/kg     | 51  | 17                        | Off ART             | B     | 47,650                    | 1129   |
| 2D3      | 30 mg/kg     | 33  | 0.5                       | ART naïve           | ND    | 640                       | 618  |
| 2E1      | 30 mg/kg     | 21  | 2                         | ART naïve           | B     | 15,780                    | 847  |
| 2E2      | 30 mg/kg     | 46  | 1.5                       | ART naïve           | B     | 6,990                     | 513  |
| 2E3      | 30 mg/kg     | 23  | 1.5                       | ART naïve           | BF    | 22,030                    | 590  |
| 2E4      | 30 mg/kg     | 38  | 1                         | ART naïve           | ND    | 32,220                    | 603  |
| 2E5      | 30 mg/kg     | 30  | 1                         | ART naïve           | ND    | 3,610                     | 532  |

**b**

| Adverse Events                 | No. AEs | % of reported AEs | No. possibly related | No. Mild | No. Moderate | No. Severe | Uninfected (No. of AEs) |         |          |          | HIV-1-infected (No. of AEs) |         |          |          |
|--------------------------------|---------|-------------------|----------------------|----------|--------------|------------|-------------------------|---------|----------|----------|-----------------------------|---------|----------|----------|
|                                |         |                   |                      |          |              |            | 1 mg/kg                 | 3 mg/kg | 10 mg/kg | 30 mg/kg | 1 mg/kg                     | 3 mg/kg | 10 mg/kg | 30 mg/kg |
| Rhinorrhea and/or cough        | 10      | 16.9              | 3                    | 8        | 2            | 0          | 1                       | 1       | 0        | 1        | 1                           | 0       | 0        | 6        |
| Malaise                        | 7       | 11.9              | 5                    | 5        | 2            | 0          | 1                       | 1       | 2        | 0        | 1                           | 2       | 0        | 0        |
| Headache                       | 6       | 10.2              | 3                    | 5        | 1            | 0          | 0                       | 0       | 0        | 1        | 1                           | 0       | 1        | 3        |
| Diarrhea                       | 5       | 8.5               | 1                    | 2        | 1            | 0          | 1                       | 1       | 0        | 0        | 1                           | 0       | 1        | 1        |
| Myalgia/arthralgia (localized) | 4       | 6.8               | 1                    | 3        | 1            | 0          | 0                       | 0       | 0        | 0        | 1                           | 0       | 3        | 0        |
| Sore throat                    | 4       | 6.8               | 2                    | 4        | 0            | 0          | 1                       | 1       | 0        | 0        | 0                           | 1       | 1        | 0        |
| Tenderness                     | 3       | 5.1               | 1                    | 3        | 0            | 0          | 0                       | 0       | 0        | 1        | 1                           | 0       | 1        | 0        |
| Increased Lacrimation          | 2       | 3.4               | 2                    | 2        | 0            | 0          | 0                       | 0       | 0        | 1        | 0                           | 0       | 0        | 1        |
| Myalgia                        | 2       | 3.4               | 1                    | 2        | 0            | 0          | 0                       | 0       | 0        | 0        | 0                           | 0       | 1        | 1        |
| Chills                         | 2       | 3.4               | 0                    | 1        | 1            | 0          | 0                       | 1       | 0        | 0        | 1                           | 0       | 0        | 0        |
| Conjunctival erythema          | 2       | 3.4               | 2                    | 2        | 0            | 0          | 0                       | 0       | 1        | 0        | 1                           | 0       | 0        | 0        |
| Fevershiness                   | 2       | 3.4               | 0                    | 1        | 1            | 0          | 0                       | 0       | 0        | 0        | 1                           | 1       | 0        | 0        |
| Nausea                         | 2       | 3.4               | 1                    | 2        | 0            | 0          | 0                       | 0       | 0        | 1        | 1                           | 0       | 0        | 0        |
| Pruritus                       | 2       | 3.4               | 2                    | 2        | 0            | 0          | 0                       | 0       | 0        | 0        | 0                           | 0       | 0        | 2        |
| Blurry vision                  | 1       | 1.7               | 1                    | 1        | 0            | 0          | 0                       | 0       | 0        | 0        | 0                           | 0       | 1        | 0        |
| Decreased appetite             | 1       | 1.7               | 0                    | 1        | 0            | 0          | 0                       | 0       | 0        | 0        | 1                           | 0       | 0        | 0        |
| Erythema                       | 1       | 1.7               | 1                    | 1        | 0            | 0          | 0                       | 0       | 0        | 1        | 0                           | 0       | 0        | 0        |
| Paresthesia upper extremity    | 1       | 1.7               | 0                    | 1        | 0            | 0          | 0                       | 0       | 0        | 0        | 1                           | 0       | 0        | 0        |
| Shingles                       | 1       | 1.7               | 0                    | 0        | 1            | 0          | 0                       | 0       | 0        | 0        | 0                           | 0       | 0        | 1        |
| Vomiting                       | 1       | 1.7               | 1                    | 1        | 0            | 0          | 0                       | 0       | 0        | 0        | 1                           | 0       | 0        | 0        |

**a.** Baseline characteristics of HIV-1-infected individuals. \*Absolute CD4 T-cell count was 309 and 302 cells  $\mu\text{l}^{-1}$  at screening. \*\*DRV/r/TDF/FTC, darunavir, ritonavir tenofovir, emtricitabine; ATV/r/3TC/ZDV, atazanavir, ritonavir, lamivudine, zidovudine. ND, not determined. **b.** 3BNC117 safety data. AE, adverse events. Subject 2D3 developed herpes zoster involving a lumbar dermatome 35 days after infusion. The event was graded as moderate and considered not related to 3BNC117.

Extended Data Table 2 | HIV-1 RNA levels and viral decay mixed-effect linear model

|          |                     |        |                   |                     |                     |        |                   |                     |                     |        |                   |                     |          |                                  |         |        |       |         |
|----------|---------------------|--------|-------------------|---------------------|---------------------|--------|-------------------|---------------------|---------------------|--------|-------------------|---------------------|----------|----------------------------------|---------|--------|-------|---------|
| 1 mg/kg  | 2A1 HIV-1 RNA Level |        |                   |                     | 2A3 HIV-1 RNA Level |        |                   |                     | 2A4 HIV-1 RNA Level |        |                   |                     | 1 mg/kg  | Contrast Dose Average SE p.value |         |        |       |         |
|          | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> |          | Screen - Day 0                   | 1mg/kg  | -0.133 | 0.258 | 0.607   |
|          | Screen              | 1,340  | 3.13              | -                   | Screen              | 41,670 | 4.62              | -                   | Screen              | 5,350  | 3.73              | -                   |          | Day -7 - Day 0                   | 1mg/kg  | -0.112 | 0.200 | 0.573   |
|          | Pre                 | 1,100  | 3.04              | -                   | Pre                 | 48,100 | 4.68              | -                   | Pre                 | 6,490  | 3.81              | -                   |          | Day 1 - Day 0                    | 1mg/kg  | 0.299  | 0.200 | 0.134   |
|          | Day 0               | 3,210  | 3.51              | 0.00                | Day 0               | 43,560 | 4.64              | 0.00                | Day 0               | 5,340  | 3.73              | 0.00                |          | Day 4 - Day 0                    | 1mg/kg  | 0.043  | 0.258 | 0.866   |
|          | Day 1               | 10,810 | 4.03              | 0.52                | Day 1               | 50,180 | 4.70              | 0.06                | Day 1               | 10,880 | 4.04              | 0.31                |          | Day 7 - Day 0                    | 1mg/kg  | -0.047 | 0.290 | 0.873   |
|          | Day 4               | 4,970  | 3.70              | 0.19                | Day 4               | 24,050 | 4.38              | -0.26               | Day 4               | 8,430  | 3.93              | 0.20                |          | Day 14 - Day 0                   | 1mg/kg  | 0.028  | 0.310 | 0.927   |
|          | Day 7               | 3,190  | 3.50              | -0.01               | Day 7               | 34,850 | 4.54              | -0.10               | Day 7               | 4,870  | 3.69              | -0.04               |          | Day 21 - Day 0                   | 1mg/kg  | 0.046  | 0.323 | 0.888   |
|          | Day 14              | 2,600  | 3.41              | -0.10               | Day 14              | 31,770 | 4.50              | -0.14               | Day 14              | 11,000 | 4.04              | 0.31                |          | Day 28 - Day 0                   | 1mg/kg  | 0.105  | 0.331 | 0.752   |
|          | Day 21              | 3,820  | 3.58              | 0.07                | Day 21              | 35,840 | 4.55              | -0.09               | Day 21              | 7,470  | 3.87              | 0.14                |          | Day 42 - Day 0                   | 1mg/kg  | 0.180  | 0.337 | 0.592   |
| 3 mg/kg  | 2B1 HIV-1 RNA Level |        |                   |                     | 2B2 HIV-1 RNA Level |        |                   |                     | 2B3 HIV-1 RNA Level |        |                   |                     | 3 mg/kg  | Contrast Dose Average SE p.value |         |        |       |         |
|          | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> |          | Screen - Day 0                   | 3mg/kg  | 0.172  | 0.316 | 0.587   |
|          | Screen              | 6,420  | 3.81              | -                   | Screen              | 80*    | 1.90              | -                   | Screen              | 53,660 | 4.73              | -                   |          | Day -7 - Day 0                   | 3mg/kg  | -0.020 | 0.244 | 0.936   |
|          | Pre                 | 4,950  | 3.69              | -                   | Pre                 | 30     | 1.48              | -                   | Pre                 | 28,810 | 4.46              | -                   |          | Day 1 - Day 0                    | 3mg/kg  | 0.061  | 0.244 | 0.803   |
|          | Day 0               | 4,090  | 3.61              | 0.00                | Day 0               | 100    | 2.00              | 0.00                | Day 0               | 38,190 | 4.58              | 0.00                |          | Day 4 - Day 0                    | 3mg/kg  | -0.351 | 0.316 | 0.267   |
|          | Day 1               | 6,470  | 3.81              | 0.20                | Day 1               | 140    | 2.15              | 0.15                | Day 1               | 31,950 | 4.50              | -0.08               |          | Day 7 - Day 0                    | 3mg/kg  | -0.338 | 0.356 | 0.343   |
|          | Day 4               | 2,690  | 3.43              | -0.18               | Day 4               | 190    | 2.28              | 0.28                | Day 4               | 11,540 | 4.06              | -0.52               |          | Day 14 - Day 0                   | 3mg/kg  | 0.024  | 0.380 | 0.949   |
|          | Day 7               | 1,610  | 3.21              | -0.40               | Day 7               | 20     | 1.30              | -0.70               | Day 7               | 20,500 | 4.31              | -0.27               |          | Day 21 - Day 0                   | 3mg/kg  | 0.022  | 0.396 | 0.955   |
|          | Day 14              | 4,880  | 3.69              | 0.08                | Day 14              | 130    | 2.11              | 0.11                | Day 14              | 35,830 | 4.55              | -0.03               |          | Day 28 - Day 0                   | 3mg/kg  | -0.063 | 0.406 | 0.877   |
|          | Day 21              | 4,160  | 3.62              | 0.01                | Day 21              | 90     | 1.95              | -0.05               | Day 21              | 41,570 | 4.62              | 0.04                |          | Day 42 - Day 0                   | 3mg/kg  | -0.017 | 0.412 | 0.966   |
| 10 mg/kg | 2C2 HIV-1 RNA Level |        |                   |                     | 2C4 HIV-1 RNA Level |        |                   |                     | 2C5 HIV-1 RNA Level |        |                   |                     | 10 mg/kg | Contrast Dose Average SE p.value |         |        |       |         |
|          | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> |          | Screen - Day 0                   | 10mg/kg | 0.415  | 0.316 | 0.189   |
|          | Screen              | 20     | 1.30              | -                   | Screen              | 2,510  | 3.41              | -                   | Screen              | 19,970 | 4.30              | -                   |          | Day -7 - Day 0                   | 10mg/kg | 0.098  | 0.244 | 0.687   |
|          | Pre                 | 30     | 1.48              | -                   | Pre                 | 1,100  | 3.04              | -                   | Pre                 | 10,860 | 4.04              | -                   |          | Day 1 - Day 0                    | 10mg/kg | -0.159 | 0.244 | 0.516   |
|          | Day 0               | 30     | 1.48              | 0.00                | Day 0               | 820    | 2.91              | 0.00                | Day 0               | 9,260  | 3.97              | 0.00                |          | Day 4 - Day 0                    | 10mg/kg | -0.260 | 0.316 | 0.410   |
|          | Day 1               | 60     | 1.78              | 0.30                | Day 1               | 690    | 2.84              | -0.07               | Day 1               | 5,300  | 3.72              | -0.25               |          | Day 7 - Day 0                    | 10mg/kg | -0.636 | 0.356 | 0.074   |
|          | Day 4               | 50     | 1.70              | 0.22                | Day 4               | 980    | 2.99              | 0.08                | Day 4               | 2,340  | 3.37              | -0.60               |          | Day 14 - Day 0                   | 10mg/kg | -0.503 | 0.380 | 0.186   |
|          | Day 7               | 9      | 0.96              | -0.52               | Day 7               | 990    | 3.00              | 0.09                | Day 7               | 410    | 2.61              | -1.36               |          | Day 21 - Day 0                   | 10mg/kg | 0.111  | 0.396 | 0.780   |
|          | Day 14              | 6      | 0.79              | -0.69               | Day 14              | 700    | 2.85              | -0.06               | Day 14              | 1,070  | 3.03              | -0.94               |          | Day 28 - Day 0                   | 10mg/kg | -0.075 | 0.406 | 0.854   |
|          | Day 21              | 20     | 1.30              | -0.18               | Day 21              | 1,420  | 3.15              | 0.24                | Day 21              | 8,900  | 3.95              | -0.02               |          | Day 42 - Day 0                   | 10mg/kg | 0.102  | 0.412 | 0.804   |
| 30 mg/kg | 2D1 HIV-1 RNA Level |        |                   |                     | 2C1 HIV-1 RNA Level |        |                   |                     | 2D3 HIV-1 RNA Level |        |                   |                     | 30 mg/kg | Contrast Dose Average SE p.value |         |        |       |         |
|          | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> |          | Screen - Day 0                   | 30mg/kg | -0.084 | 0.158 | 0.593   |
|          | Screen              | 5,730  | 3.76              | -                   | Screen              | 31,870 | 4.50              | -                   | Screen              | 3,620  | 3.56              | -                   |          | Day -7 - Day 0                   | 30mg/kg | 0.022  | 0.122 | 0.860   |
|          | Pre                 | 34,840 | 4.54              | -                   | Pre                 | 31,870 | 4.50              | -                   | Pre                 | 3,280  | 3.52              | -                   |          | Day 1 - Day 0                    | 30mg/kg | 0.065  | 0.122 | 0.593   |
|          | Day 0               | 53,470 | 4.73              | 0.00                | Day 0               | 47,650 | 4.68              | 0.00                | Day 0               | 640    | 2.81              | 0.00                |          | Day 4 - Day 0                    | 30mg/kg | -0.458 | 0.158 | 0.004   |
|          | Day 1               | 84,450 | 4.93              | 0.20                | Day 1               | 46,040 | 4.66              | -0.02               | Day 1               | 450    | 2.65              | -0.16               |          | Day 7 - Day 0                    | 30mg/kg | -1.043 | 0.178 | < 0.001 |
|          | Day 4               | 36,140 | 4.56              | -0.17               | Day 4               | 7,790  | 3.89              | -0.79               | Day 4               | 210    | 2.32              | -0.49               |          | Day 14 - Day 0                   | 30mg/kg | -1.181 | 0.190 | < 0.001 |
|          | Day 7               | 5,980  | 3.78              | -0.95               | Day 7               | 7,470  | 3.87              | -0.81               | Day 7               | 80     | 1.90              | -0.91               |          | Day 21 - Day 0                   | 30mg/kg | -0.866 | 0.198 | < 0.001 |
|          | Day 14              | 4,830  | 3.68              | -1.05               | Day 14              | 19,360 | 4.29              | -0.39               | Day 14              | 50     | 1.70              | -1.11               |          | Day 28 - Day 0                   | 30mg/kg | -0.515 | 0.203 | 0.011   |
|          | Day 21              | 15,180 | 4.18              | -0.55               | Day 21              | 30,380 | 4.48              | -0.20               | Day 21              | 3      | 0.52              | -1.77               |          | Day 42 - Day 0                   | 30mg/kg | -0.339 | 0.206 | 0.099   |
| 30 mg/kg | 2E1 HIV-1 RNA Level |        |                   |                     | 2E2 HIV-1 RNA Level |        |                   |                     | 2E3 HIV-1 RNA Level |        |                   |                     | 30 mg/kg | Contrast Dose Average SE p.value |         |        |       |         |
|          | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> |          | Screen - Day 0                   | 30mg/kg | -0.371 | 0.213 | 0.081   |
|          | Screen              | 8,742  | 3.94              | -                   | Screen              | 3,394  | 3.53              | -                   | Screen              | 23,168 | 4.36              | -                   |          | Day 1 - Day 0                    | 30mg/kg | -0.371 | 0.213 | 0.081   |
|          | Pre                 | 12,031 | 4.08              | -                   | Pre                 | 1,802  | 3.26              | -                   | Pre                 | 30,570 | 4.49              | -                   |          | Day 4 - Day 0                    | 30mg/kg | -0.458 | 0.158 | 0.004   |
|          | Day 0               | 15,780 | 4.20              | 0.00                | Day 0               | 6,990  | 3.84              | 0.00                | Day 0               | 22,030 | 4.34              | 0.00                |          | Day 7 - Day 0                    | 30mg/kg | -1.043 | 0.178 | < 0.001 |
|          | Day 1               | 14,790 | 4.17              | -0.03               | Day 1               | 6,450  | 3.81              | -0.03               | Day 1               | 38,620 | 4.59              | 0.25                |          | Day 14 - Day 0                   | 30mg/kg | -1.181 | 0.190 | < 0.001 |
|          | Day 4               | 3,560  | 3.55              | -0.65               | Day 4               | 2,340  | 3.37              | -0.47               | Day 4               | 10,540 | 4.02              | -0.32               |          | Day 21 - Day 0                   | 30mg/kg | -0.866 | 0.198 | < 0.001 |
|          | Day 7               | 404    | 2.61              | -1.59               | Day 7               | 1,663  | 3.22              | -0.62               | Day 7               | 1,308  | 3.12              | -1.22               |          | Day 28 - Day 0                   | 30mg/kg | -0.515 | 0.203 | 0.011   |
|          | Day 14              | 468    | 2.67              | -1.53               | Day 14              | 253    | 2.40              | -1.44               | Day 14              | 3,901  | 3.59              | -0.75               |          | Day 42 - Day 0                   | 30mg/kg | -0.339 | 0.206 | 0.099   |
|          | Day 21              | 8,557  | 3.93              | -0.27               | Day 21              | 917    | 2.96              | -0.88               | Day 21              | 23,717 | 4.38              | 0.04                |          | Day 56 - Day 0                   | 30mg/kg | -0.371 | 0.213 | 0.081   |
| 30 mg/kg | 2E4 HIV-1 RNA Level |        |                   |                     | 2E5 HIV-1 RNA Level |        |                   |                     | 2E6 HIV-1 RNA Level |        |                   |                     | 30 mg/kg | Contrast Dose Average SE p.value |         |        |       |         |
|          | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> | Timepoint           | c/ml   | log <sub>10</sub> | Δ log <sub>10</sub> |          | Screen - Day 0                   | 30mg/kg | -0.371 | 0.213 | 0.081   |
|          | Screen              | 45,311 | 4.66              | -                   | Screen              | 4,720  | 3.67              | -                   | Screen              | 27,998 | 4.45              | 0.10                |          | Day 1 - Day 0                    | 30mg/kg | -0.371 | 0.213 | 0.081   |
|          | Pre                 | 66,889 | 4.83              | -                   | Pre                 | 4,244  | 3.63              | -                   | Pre                 | 3,280  | 3.52              | -                   |          | Day 4 - Day 0                    | 30mg/kg | -0.458 | 0.158 | 0.004   |
|          | Day 0               | 32,220 | 4.51              | 0.00                | Day 0               | 3,610  | 3.56              | 0.00                | Day 0               | 640    | 2.81              | 0.00                |          | Day 7 - Day 0                    | 30mg/kg | -1.043 | 0.178 | < 0.001 |
|          | Day 1               | 44,980 | 4.65              | 0.14                | Day 1               | 5,300  | 3.72              | 0.16                | Day 1               | 38,620 | 4.59              | 0.25                |          | Day 14 - Day 0                   | 30mg/kg | -1.181 | 0.190 | < 0.001 |
|          | Day 4               | 19,290 | 4.29              | -0.22               | Day 4               | 1,000  | 3.00              | -0.56               | Day 4               | 10,540 | 4.02              | -0.32               |          | Day 21 - Day 0                   | 30mg/kg | -0.866 | 0.198 | < 0.001 |
|          | Day 7               | 3,642  | 3.56              | -0.95               | Day 7               | 181    | 2.26              | -1.30               | Day 7               | 1,308  | 3.12              | -1.22               |          | Day 28 - Day 0                   | 30mg/kg | -0.515 | 0.203 | 0.011   |
|          | Day 14              | 983    | 2.99              | -1.52               | Day 14              | 77     | 1.89              | -1.67               | Day 14              | 3,901  | 3.59              | -0.75               |          | Day 42 - Day 0                   | 30mg/kg | -0.339 | 0.206 | 0.099   |
|          | Day 21              | 5,726  | 3.76              | -0.75               | Day 21              | 11     | 1.04              | -2.52               | Day 21              | 23,717 | 4.38              | 0.04                |          | Day 56 - Day 0                   | 30mg/kg | -0.371 | 0.213 | 0.081   |

a. HIV-1 RNA levels. Subjects 2B2 and 2C2 were on ART. Subject 2D3 started ART after day 42. Screen was performed between day -49 and day -14. Viraemia measurements at "Pre" were performed between day -42 and day -2. b. Viral decay mixed-effect linear model.

Extended Data Table 3 | Sensitivity of autologous virus isolates and cloned HIV-1 envelopes to 3BNC117

| Autologous virus isolates |     |                   |                                  |                          |
|---------------------------|-----|-------------------|----------------------------------|--------------------------|
| Dose                      | ID  | Day post infusion | 3BNC117 IC <sub>50</sub> (µg/ml) | 3BNC117 IC <sub>50</sub> |
| 1 mg/kg                   | 2A1 | Day 0             | ND                               | ND                       |
|                           |     | Day 28            | 0.90                             |                          |
|                           | 2A3 | Day 0             | 0.11                             | 35.3                     |
|                           |     | Day 28            | 3.78                             |                          |
|                           | 2A4 | Day 0             | 0.07                             | 13.5                     |
|                           |     | Day 28            | 0.94                             |                          |
| 3 mg/kg                   | 2B1 | Day 0             | 0.77                             | 25.9                     |
|                           |     | Day 28            | >20                              |                          |
|                           | 2B2 | Day 0             | >20                              | ND                       |
|                           |     | Day 28            | ND                               |                          |
|                           | 2B3 | Day 0             | 0.20                             | 1.5                      |
|                           |     | Day 28            | 0.30                             |                          |
| 10 mg/kg                  | 2C2 | Day 0             | 0.49                             | 0.1                      |
|                           |     | Day 56            | 0.03                             |                          |
|                           | 2C4 | Day 0             | >20                              | No change                |
|                           |     | Day 28            | >20                              |                          |
|                           | 2C5 | Day 0             | 0.09                             | 167.0                    |
|                           |     | Day 28            | 15.36                            |                          |
| 30 mg/kg                  | 2C1 | Day -7            | 0.54                             | 1.8                      |
|                           |     | Day 28            | 1.78                             |                          |
|                           | 2D1 | Day 0             | 0.68                             | 1.3                      |
|                           |     | Day 28            | 0.90                             |                          |
|                           | 2D3 | Day 0             | 0.13                             | 2.7                      |
|                           |     | Day 28            | 0.35                             |                          |
|                           | 2E1 | Screen            | 0.40                             | ND                       |
|                           |     | Day 28            | ND                               |                          |
|                           | 2E2 | Screen            | 0.18                             | 12.3                     |
|                           |     | Day 28            | 2.23                             |                          |
|                           | 2E3 | Screen            | 0.18                             | 6.1                      |
|                           |     | Day 28            | 1.10                             |                          |
|                           | 2E4 | Screen            | 0.24                             | ND                       |
|                           |     | Day 28            | ND                               |                          |
|                           | 2E5 | Screen            | 0.30                             | ND                       |
|                           |     | Day 28            | ND                               |                          |

| HIV-1 envelopes cloned from plasma |                   |                 |                                    |                     |
|------------------------------------|-------------------|-----------------|------------------------------------|---------------------|
| Clone                              | Cloning procedure | Vector backbone | 3BNC117 (IC <sub>50</sub> ; µg/ml) | Average (geo. mean) |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| 2C5_D0_12                          | gp120             | pSVIII          | 0.015                              | 0.02                |
| 2C5_D0_21                          | gp120             | pSVIII          | 0.017                              |                     |
| 2C5_D0_27                          | gp120             | pSVIII          | 0.017                              |                     |
| 2C5_W4_59                          | gp120             | pSVIII          | 11.543                             | 7.09                |
| 2C5_W4_22                          | gp120             | pSVIII          | 6.737                              |                     |
| 2C5_W4_27                          | gp120             | pSVIII          | 7.514                              |                     |
| 2C5_W4_28                          | gp120             | pSVIII          | 3.495                              |                     |
| 2C5_W4_34                          | gp120             | pSVIII          | 8.758                              |                     |
| 2C1_D0_12                          | gp120             | pSVIII          | 0.209                              | 0.06                |
| 2C1_D0_22                          | gp120             | pSVIII          | 0.158                              |                     |
| 2C1_D0_32                          | gp120             | pSVIII          | 0.006                              |                     |
| 2C1_W4_12                          | gp120             | pSVIII          | 0.043                              | 0.14                |
| 2C1_W4_18                          | gp120             | pSVIII          | 0.225                              |                     |
| 2C1_W4_31                          | gp120             | pSVIII          | 0.262                              |                     |
| 2D1_D0_D5                          | gp160             | pcDNA3.1        | 0.165                              | 0.15                |
| 2D1_D0_B3.1                        | gp160             | pcDNA3.1        | 0.128                              |                     |
| 2D1_D0_B10                         | gp160             | pcDNA3.1        | 0.172                              |                     |
| 2D1_W4_37                          | gp120             | pSVIII          | 0.578                              | 0.52                |
| 2D1_W4_40                          | gp120             | pSVIII          | 0.501                              |                     |
| 2D1_W4_69                          | gp120             | pSVIII          | 0.465                              |                     |
| 2D1_W4_71                          | gp120             | pSVIII          | 0.523                              |                     |
| ND                                 |                   |                 |                                    |                     |
| 2E1_D0_12                          | gp160             | pcDNA3.1        | 0.103                              | 0.09                |
| 2E1_D0_20                          | gp160             | pcDNA3.1        | 0.115                              |                     |
| 2E1_D0_34                          | gp160             | pcDNA3.1        | 0.068                              |                     |
| 2E1_W4_23                          | gp160             | pcDNA3.1        | 0.041                              | 0.23                |
| 2E1_W4_E1                          | gp160             | pcDNA3.1        | 0.590                              |                     |
| 2E1_W4_F6                          | gp160             | pcDNA3.1        | 0.496                              |                     |
| 2E2_D0_A10                         | gp160             | pcDNA3.1        | 0.017                              | 0.01                |
| 2E2_D0_C3                          | gp160             | pcDNA3.1        | 0.010                              |                     |
| 2E2_D0_E9                          | gp160             | pcDNA3.1        | 0.018                              |                     |
| 2E2_W4_B9                          | gp160             | pcDNA3.1        | 0.017                              | 0.03                |
| 2E2_W4_C11                         | gp160             | pcDNA3.1        | 0.020                              |                     |
| 2E2_W4_D5                          | gp160             | pcDNA3.1        | 0.057                              |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |
| ND                                 |                   |                 |                                    |                     |

ND, not determined.

Extended Data Table 4 | Pharmacokinetics of 3BNC117 based on a 56-day period post infusion

| ID  | HIV-1/ART    | 3BNC117<br>(mg/kg) | Method | Cmax<br>(µg/ml) | adjusted<br>R-squared | Estimated<br>T <sub>1/2</sub> (days) | Lambda<br>(lower)* | Lambda<br>(upper)* | AUC<br>(INF_pred) | Tlast* | Clast<br>(µg/ml) | AUC_%<br>Extrap_pred |
|-----|--------------|--------------------|--------|-----------------|-----------------------|--------------------------------------|--------------------|--------------------|-------------------|--------|------------------|----------------------|
| 1A1 | Negative     | 1                  | ELISA  | 27.4            | n.d.                  | n.d.                                 | -                  | -                  | n.d.              | -      | -                | n.d.                 |
|     |              |                    | TZM.bl | 27.4            | 0.936                 | 21.32                                | 14                 | 42                 | 184.0             | 42     | 1.2              | 19.2                 |
| 1A2 | Negative     | 1                  | ELISA  | 18.8            | n.d.                  | n.d.                                 | -                  | -                  | n.d.              | -      | -                | n.d.                 |
|     |              |                    | TZM.bl | 27.8            | 0.893                 | 12.89                                | 7                  | 42                 | 181.7             | 42     | 1.0              | 9.0                  |
| 1A3 | Negative     | 1                  | ELISA  | 11.2            | n.d.                  | n.d.                                 | -                  | -                  | n.d.              | -      | -                | n.d.                 |
|     |              |                    | TZM.bl | 23.5            | 0.976                 | 10.38                                | 4                  | 28                 | 138.0             | 28     | 1.3              | 14.0                 |
| 2A1 | Positive/Off | 1                  | ELISA  | 15.7            | n.d.                  | n.d.                                 | -                  | -                  | n.d.              | -      | -                | n.d.                 |
| 2A3 | Positive/Off | 1                  | ELISA  | 22.7            | n.d.                  | n.d.                                 | -                  | -                  | n.d.              | -      | -                | n.d.                 |
| 2A4 | Positive/Off | 1                  | ELISA  | 33.8            | n.d.                  | n.d.                                 | -                  | -                  | n.d.              | -      | -                | n.d.                 |
|     |              |                    | TZM.bl | 15.6            | n.d.                  | n.d.                                 | -                  | -                  | n.d.              | -      | -                | n.d.                 |
| 1B1 | Negative     | 3                  | ELISA  | 89.9            | 0.985                 | 19.64                                | 7                  | 56                 | 685.2             | 56     | 3.1              | 12.1                 |
|     |              |                    | TZM.bl | 70.9            | 0.881                 | 24.30                                | 7                  | 56                 | 483.7             | 56     | 2.7              | 17.9                 |
| 1B2 | Negative     | 3                  | ELISA  | 90.1            | 0.878                 | 12.91                                | 7                  | 42                 | 523.5             | 42     | 2.2              | 8.8                  |
|     |              |                    | TZM.bl | 116.0           | 0.936                 | 10.04                                | 7                  | 42                 | 481.2             | 42     | 1.4              | 4.6                  |
| 1B3 | Negative     | 3                  | ELISA  | 243.4           | 0.954                 | 21.54                                | 14                 | 56                 | 1017.3            | 56     | 4.2              | 12.1                 |
|     |              |                    | TZM.bl | 74.5            | 0.954                 | 21.70                                | 7                  | 56                 | 706.6             | 56     | 3.9              | 15.6                 |
| 2B1 | Positive/Off | 3                  | ELISA  | 90.8            | 0.942                 | 9.60                                 | 4                  | 21                 | 219.4             | 21     | 2.8              | 19.2                 |
|     |              |                    | TZM.bl | 57.4            | 0.994                 | 10.28                                | 7                  | 28                 | 245.9             | 28     | 1.9              | 12.1                 |
| 2B2 | Positive/On  | 3                  | ELISA  | 97.6            | 0.928                 | 8.74                                 | 7                  | 21                 | 200.7             | 21     | 2.5              | 16.5                 |
| 2B3 | Positive/Off | 3                  | ELISA  | 60.0            | 0.981                 | 9.08                                 | 4                  | 21                 | 251.2             | 21     | 3.5              | 17.0                 |
|     |              |                    | TZM.bl | 32.6            | 0.994                 | 9.19                                 | 4                  | 28                 | 129.8             | 28     | 1.0              | 10.1                 |
| 1C1 | Negative     | 10                 | ELISA  | 347.8           | 0.942                 | 16.33                                | 7                  | 56                 | 1667.9            | 56     | 5.9              | 7.5                  |
|     |              |                    | TZM.bl | 281.8           | 0.964                 | 15.79                                | 7                  | 56                 | 1598.9            | 56     | 5.6              | 7.3                  |
| 1C3 | Negative     | 10                 | ELISA  | 308.9           | 0.860                 | 29.01                                | 7                  | 56                 | 2529.4            | 56     | 15.8             | 23.5                 |
|     |              |                    | TZM.bl | 244.7           | 0.700                 | 26.59                                | 14                 | 56                 | 2228.5            | 56     | 13.1             | 19.3                 |
| 1C4 | Negative     | 10                 | ELISA  | 180.6           | 0.835                 | 10.73                                | 7                  | 42                 | 859.1             | 42     | 4.0              | 5.5                  |
|     |              |                    | TZM.bl | 230.7           | 0.990                 | 10.56                                | 14                 | 56                 | 1022.0            | 56     | 1.4              | 2.2                  |
| 2C2 | Positive/On  | 10                 | ELISA  | 326.8           | 0.866                 | 10.87                                | 7                  | 42                 | 904.5             | 42     | 2.7              | 5.6                  |
| 2C4 | Positive/Off | 10                 | ELISA  | 178.2           | 0.944                 | 6.92                                 | 4                  | 28                 | 550.2             | 28     | 2.8              | 5.1                  |
|     |              |                    | TZM.bl | 193.0           | 0.979                 | 6.14                                 | 4                  | 28                 | 417.9             | 28     | 1.8              | 3.4                  |
| 2C5 | Positive/Off | 10                 | ELISA  | 201.1           | 0.997                 | 6.60                                 | 7                  | 28                 | 592.6             | 28     | 2.6              | 4.1                  |
|     |              |                    | TZM.bl | 174.4           | 0.994                 | 6.50                                 | 7                  | 28                 | 431.2             | 28     | 1.8              | 3.7                  |
| 1E2 | Negative     | 30                 | ELISA  | 360.8           | 0.930                 | 14.48                                | 7                  | 56                 | 4259.4            | 56     | 13.6             | 5.9                  |
|     |              |                    | TZM.bl | 1166.3          | 0.978                 | 16.19                                | 14                 | 56                 | 5901.2            | 56     | 16.1             | 6.2                  |
| 1E3 | Negative     | 30                 | ELISA  | 361.8           | 0.990                 | 16.26                                | 7                  | 56                 | 3177.5            | 56     | 11.7             | 8.2                  |
|     |              |                    | TZM.bl | 606.2           | 0.989                 | 17.85                                | 7                  | 56                 | 3422.2            | 56     | 12.3             | 9.3                  |
| 1E5 | Negative     | 30                 | ELISA  | 765.0           | 0.992                 | 14.15                                | 4                  | 56                 | 5874.2            | 56     | 16.3             | 5.7                  |
|     |              |                    | TZM.bl | 939.6           | 0.959                 | 15.22                                | 14                 | 56                 | 6424.5            | 56     | 19.1             | 6.1                  |
| 2C1 | Positive/Off | 30                 | ELISA  | 410.2           | 0.991                 | 5.80                                 | 7                  | 28                 | 1707.8            | 28     | 5.1              | 2.6                  |
|     |              |                    | TZM.bl | 717.4           | 0.993                 | 5.99                                 | 7                  | 42                 | 2186.4            | 42     | 1.6              | 0.6                  |
| 2D1 | Positive/Off | 30                 | ELISA  | 976.4           | 0.996                 | 6.86                                 | 7                  | 42                 | 2494.9            | 42     | 3.4              | 1.3                  |
|     |              |                    | TZM.bl | 849.3           | 0.990                 | 8.83                                 | 7                  | 56                 | 1825.5            | 56     | 1.1              | 0.8                  |
| 2D3 | Positive/Off | 30                 | ELISA  | 571.0           | 0.962                 | 13.39                                | 7                  | 56                 | 3616.5            | 56     | 9.9              | 4.8                  |
|     |              |                    | TZM.bl | 953.7           | 0.993                 | 11.23                                | 7                  | 42                 | 4346.2            | 42     | 16.0             | 6.0                  |
| 2E1 | Positive/Off | 30                 | ELISA  | 712.4           | 0.970                 | 11.14                                | 14                 | 56                 | 3028.5            | 56     | 4.6              | 2.3                  |
| 2E2 | Positive/Off | 30                 | ELISA  | 789.4           | 0.920                 | 11.14                                | 7                  | 56                 | 3596.1            | 56     | 6.7              | 2.4                  |
| 2E3 | Positive/Off | 30                 | ELISA  | 559.6           | 0.821                 | 8.54                                 | 7                  | 28                 | 2495.6            | 28     | 13.7             | 8.4                  |

Estimation of PK parameters retrieved from study participants (ID) up to 8 weeks post 3BNC117 infusion. 3BNC117 was administered at 1, 3, 10, and 30 mg kg<sup>-1</sup> in uninfected (Negative) and HIV-1-infected (Positive) individuals. 3BNC117 serum concentrations were determined by ELISA or TZM.bl assay (see Methods). PK-parameters were obtained by performing a non-compartmental analysis (NCA) using WinNonlin 6.3. T<sub>1/2</sub> was estimated using values between Lambda upper/lower. n.d., PK parameters were not determined because of insufficient data or an extrapolated AUC<sub>inf</sub> that exceeded 25%. Values in red indicate extrapolated AUC<sub>inf</sub> >10%. \*Days post infusion of 3BNC117.



Extended Data Table 5 | Summary of 3BNC117 pharmacokinetics based on a 56-day period post infusion

| Dose  | HIV-1-status | Subjects | Method<br>(subjects analyzed) | C <sub>max</sub> (µg/ml) |       |                | t <sub>1/2</sub> (days) (1) |      |             |
|-------|--------------|----------|-------------------------------|--------------------------|-------|----------------|-----------------------------|------|-------------|
|       |              |          |                               | Mean                     | SD    | Range          | Mean                        | SD   | Range       |
| 1 mg  | Neg.         | 3        | ELISA (3)                     | <b>19.1</b>              | 8.1   | 11.2 - 27.4    | <b>n.d.</b>                 | n.d. | n.d.        |
|       |              |          | TZM.bl (3)                    | <b>26.2</b>              | 2.4   | 23.5 - 27.8    | <b>n.d.</b>                 | n.d. | n.d.        |
| 1 mg  | Pos.         | 3        | ELISA (3)                     | <b>24.1</b>              | 9.1   | 15.7 - 33.8    | <b>n.d.</b>                 | n.d. | n.d.        |
|       |              |          | TZM.bl (1)                    | <b>23.5</b>              | n.d.  | n.d.           | <b>n.d.</b>                 | n.d. | n.d.        |
| 3 mg  | Neg.         | 3        | ELISA (3)                     | <b>141.1</b>             | 88.6  | 89.9 - 243.4   | <b>18.0</b>                 | 4.5  | 12.9 - 21.5 |
|       |              |          | TZM.bl (3)                    | <b>87.1</b>              | 25.1  | 70.9 - 116.0   | <b>18.7</b>                 | 7.6  | 10.0 - 24.3 |
| 3 mg  | Pos.         | 3        | ELISA (3)                     | <b>73.7</b>              | 35.7  | 32.6 - 97.6    | <b>9.2</b>                  | 0.4  | 8.7 - 9.6   |
|       |              |          | TZM.bl (3)                    | <b>50.6</b>              | 15.7  | 32.6 - 61.7    | <b>11.2</b>                 | 2.2  | 9.5 - 13.73 |
| 10 mg | Neg.         | 3        | ELISA (3)                     | <b>279.1</b>             | 87.5  | 180.6 - 347.8  | <b>18.7</b>                 | 9.4  | 10.7 - 29.0 |
|       |              |          | TZM.bl (3)                    | <b>252.4</b>             | 26.4  | 230.7 - 281.8  | <b>17.7</b>                 | 8.1  | 10.7 - 26.6 |
| 10 mg | Pos.         | 3        | ELISA (3)                     | <b>235.4</b>             | 80.0  | 178.2 - 326.8  | <b>8.1</b>                  | 2.4  | 6.6 - 10.9  |
|       |              |          | TZM.bl (3)                    | <b>213.6</b>             | 52.6  | 174.4 - 273.4  | <b>8.8</b>                  | 4.3  | 6.5 - 13.7  |
| 30 mg | Neg.         | 3        | ELISA (3)                     | <b>495.9</b>             | 233.1 | 360.8 - 765.0  | <b>15.0</b>                 | 1.1  | 14.2 - 16.3 |
|       |              |          | TZM.bl (3)                    | <b>904.0</b>             | 281.7 | 1166.3 - 606.2 | <b>16.4</b>                 | 1.3  | 15.2 - 17.9 |
| 30 mg | Pos.         | 6        | ELISA (6)                     | <b>669.8</b>             | 199.7 | 410.2 - 976.4  | <b>9.9</b>                  | 3.3  | 5.8 - 13.7  |
|       |              |          | TZM.bl (3)                    | <b>840.1</b>             | 118.4 | 717.4 - 953.7  | <b>8.9</b>                  | 2.3  | 6.7 - 11.2  |
| All   | Neg.         | 12       | ELISA (9)                     | -                        | -     | -              | <b>17.2</b>                 | 5.5  | 10.7 - 29.0 |
|       |              |          | TZM.bl (9)                    | -                        | -     | -              | <b>17.6</b>                 | 5.7  | 10.0 - 26.6 |
| All   | Pos.         | 15       | ELISA (12)                    | -                        | -     | -              | <b>9.3</b>                  | 2.6  | 5.7 - 13.7  |
|       |              |          | TZM.bl (9)                    | -                        | -     | -              | <b>9.6</b>                  | 2.9  | 6.1 - 13.7  |

(1) Estimation of half-lives; SD, standard deviation.

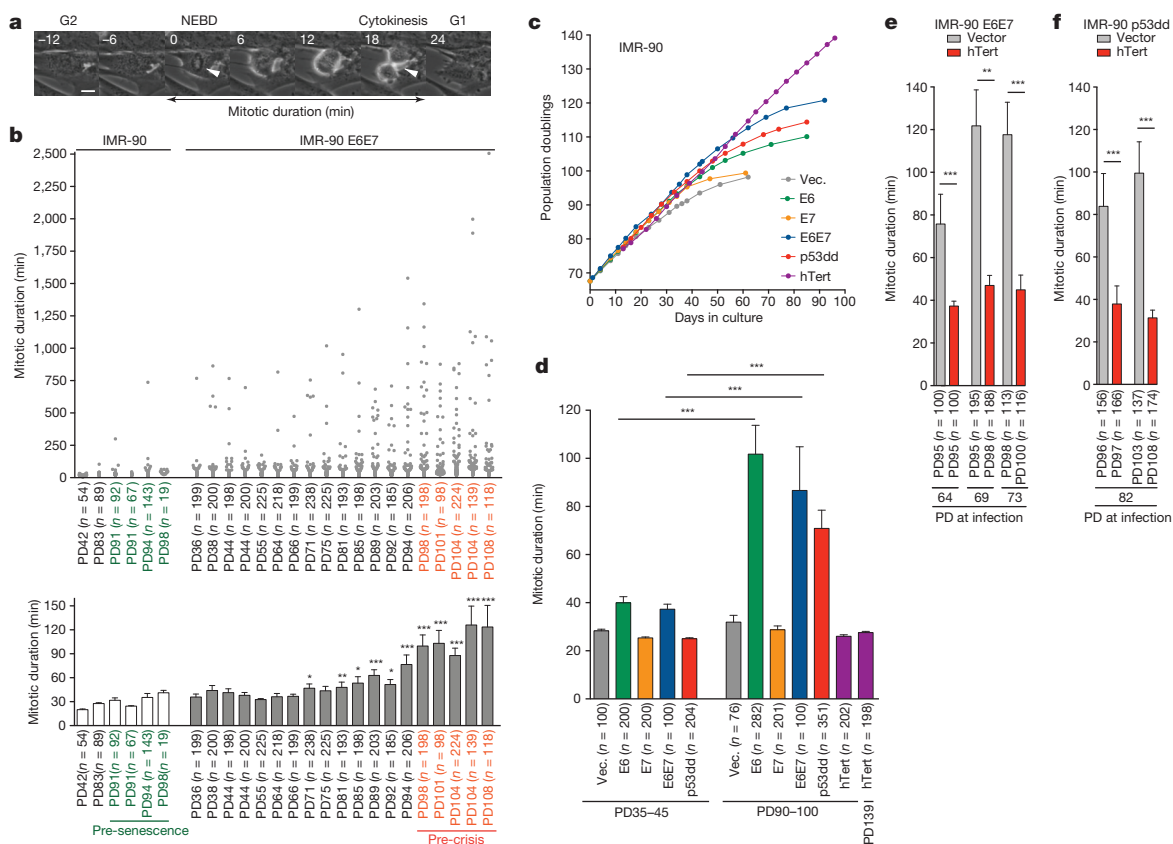
# Cell death during crisis is mediated by mitotic telomere deprotection

Makoto T. Hayashi<sup>1,2</sup>, Anthony J. Cesare<sup>1,3</sup>, Teresa Rivera<sup>1</sup> & Jan Karlseder<sup>1</sup>

Tumour formation is blocked by two barriers: replicative senescence and crisis<sup>1</sup>. Senescence is triggered by short telomeres and is bypassed by disruption of tumour-suppressive pathways. After senescence bypass, cells undergo crisis, during which almost all of the cells in the population die. Cells that escape crisis harbour unstable genomes and other parameters of transformation. The mechanism of cell death during crisis remains unexplained. Here we show that human cells in crisis undergo spontaneous mitotic arrest, resulting in death during mitosis or in the following cell cycle. This phenotype is induced by loss of p53 function, and is suppressed by telomerase overexpression. Telomere fusions triggered mitotic arrest in p53-compromised non-crisis cells, indicating that such fusions are the underlying cause of cell death.

Exacerbation of mitotic telomere deprotection by partial *TRF2* (also known as *TERF2*) knockdown<sup>2</sup> increased the ratio of cells that died during mitotic arrest and sensitized cancer cells to mitotic poisons. We propose a crisis pathway wherein chromosome fusions induce mitotic arrest, resulting in mitotic telomere deprotection and cell death, thereby eliminating precancerous cells from the population.

Replicative senescence is induced by partially deprotected telomeres, which activate a DNA damage response (DDR) without telomere fusions<sup>2</sup>. Crisis requires the bypass of senescence through loss of checkpoints and causes massive cell death concomitant with further telomere shortening and spontaneous telomere fusions. However, the mechanism of cell death was not understood. Mitotic arrest is



**Figure 1 | Spontaneous mitotic arrest upon bypass of senescence.**

**a**, Representative movie frames of normal mitosis. Arrows, nuclear envelope breakdown (NEBD) and cytokinesis. Scale bar, 10  $\mu$ m. **b**, Mitotic duration of individual indicated cells (upper panel). Mean mitotic duration  $\pm$  s.e.m. (lower panel) (Mann–Whitney test, compared to IMR-90 E6E7 PD36). **c**, Growth

curves of indicated IMR-90 derivatives. Vec., vector. **d**, Mitotic duration of indicated derivatives as in **b**. Results were reproduced independently at least twice. **e**, **f**, Mitotic duration of IMR-90 E6E7 (**e**) or p53dd (**f**) cells expressing hTert as in **b**. Cells were infected at the indicated PD. \* $P < 0.05$ , \*\* $P < 0.005$ , \*\*\* $P < 0.0001$ . Mann–Whitney test.

<sup>1</sup>The Salk Institute for Biological Studies, Molecular and Cell Biology Department, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. <sup>2</sup>Department of Gene Mechanisms, Graduate School of Biostudies/The Hakubi Center for Advanced Research, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. <sup>3</sup>Children's Medical Research Institute, University of Sydney, 214 Hawkesbury Road, Westmead, New South Wales 2145, Australia.

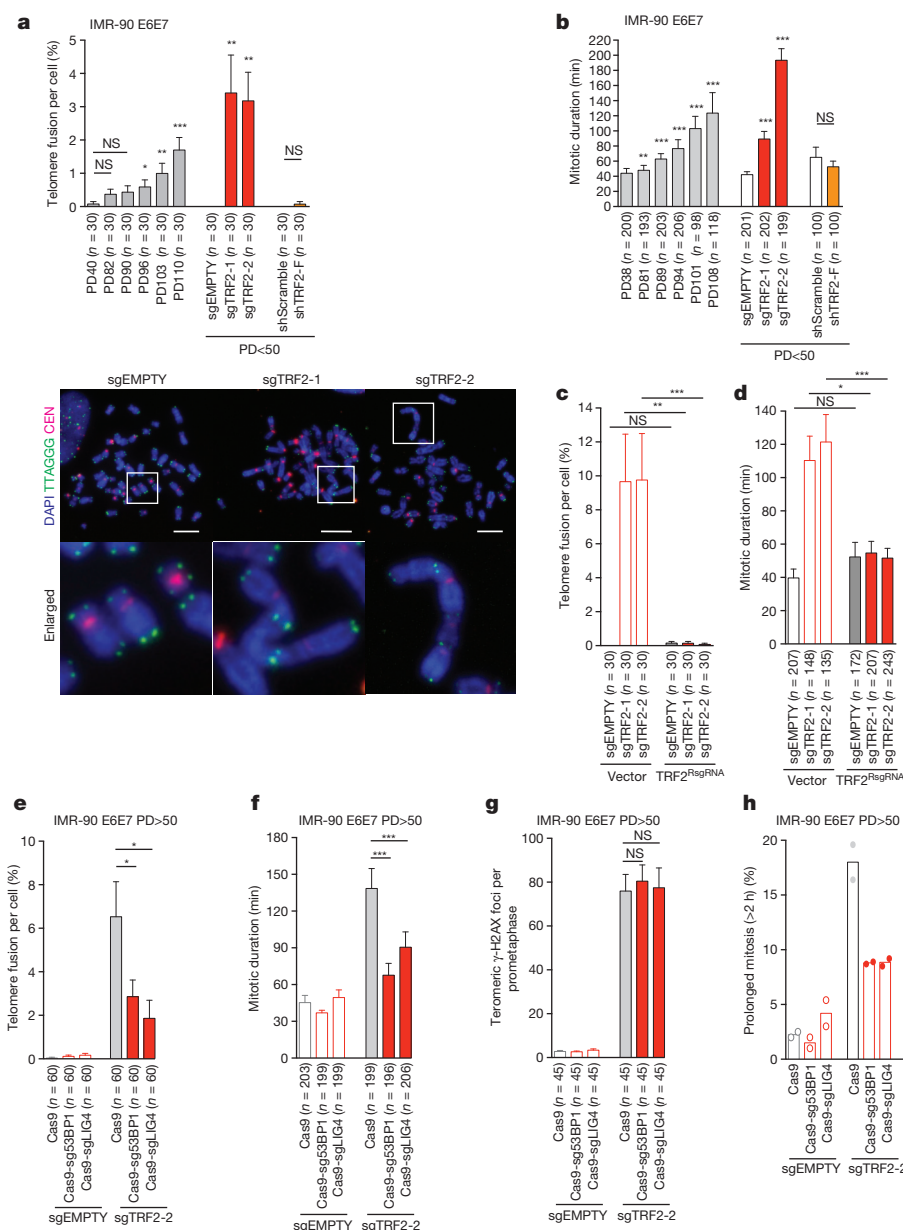
associated with spindle assembly checkpoint (SAC)-independent telomere deprotection<sup>3</sup>, and we therefore set out to test whether prolonged mitosis has a role in crisis.

We monitored mitotic duration using live-cell imaging. Mitosis in primary IMR-90 fibroblasts lasted <45 min. However, IMR-90 fibroblasts expressing HPV16 E6 and E7, which inhibit p53 and Rb<sup>4</sup>, displayed variable mitotic duration upon senescence bypass (Fig. 1a, b). Prolonged mitosis, defined as mitosis of >2 h, became prominent in pre-crisis cells (Extended Data Fig. 1a). Spontaneous mitotic arrest also occurred in pre-crisis cells following expression of E6 or dominant-negative p53dd (Extended Data Fig. 1b), indicating that loss of p53 function was required (Fig. 1c, d and Extended Data Fig. 1c, d).

Overexpressing hTERT<sup>5,6</sup> prevented senescence in IMR-90 cells (Fig. 1c, d and Extended Data Fig. 1c, d). Telomere elongation in

IMR-90 E6E7 or p53dd cells also suppressed mitotic arrest (Fig. 1e, f and Extended Data Fig. 2a–c), confirming telomere shortening as the cause. Reversine inhibition of MPS1 (ref. 7) suppressed mitotic arrest (Extended Data Fig. 1e), indicating dependence on the SAC. Hesperadin, an Aurora B kinase inhibitor required for activation of the SAC upon tensionless kinetochore–microtubule attachment<sup>8</sup>, suppressed mitotic arrest (Extended Data Fig. 1e), suggesting abnormal kinetochore–microtubule attachment.

To determine if telomere fusion causes mitotic arrest, we used two independent guide RNAs (sgTRF2-1 and sgTRF2-2)<sup>9</sup>, which efficiently reduced TRF2 expression and induced telomere fusions in young IMR-90 E6E7 cells (Fig. 2a and Extended Data Fig. 2d, e). These guide RNAs also led to mitotic arrest, comparable to IMR-90 E6E7 cells around population doubling (PD) 108 (Fig. 2b and Extended Data



**Figure 2 | Telomere fusions induce mitotic arrest.** **a**, Mean percentage of telomeric fusions per cell  $\pm$  s.e.m. in IMR-90 E6E7 and derivatives 7 days post infection around PD45. Representative metaphases below. Scale bar, 10  $\mu$ m. Enlarged images are 4 $\times$  magnified. **b**, Mean mitotic duration  $\pm$  s.e.m. in the cell lines from **a**. Mitotic duration data of growing IMR-90 E6E7 cells are the same as Fig. 1b. **c**, **d**, Percentage of telomeric fusion (**c**) and mitotic duration

(**d**) in indicated IMR-90 E6E7 cells (mean  $\pm$  s.e.m.). **e–g**, Percentage of telomeric fusion (**e**), mitotic duration (**f**) and mean number of meta-TIF (**g**) in indicated cells (mean  $\pm$  s.e.m.). **h**, Scatter plots with bars showing the mean percentage of cells with prolonged mitosis (two independent experiments). \* $P$  < 0.05, \*\* $P$  < 0.005, \*\*\* $P$  < 0.0001. NS, not significant. Mann–Whitney test.

Fig. 2f). Suppression of both telomere fusion and mitotic arrest by a resistant TRF2 (TRF2<sup>RsgRNA</sup>) excludes off-target effects (Fig. 2c, d and Extended Data Fig. 3a–c).

To address whether telomeric DDR or telomere fusion induces mitotic arrest, we deleted *TRF2* in young IMR-90 E6E7 cells lacking 53BP1 or ligase IV (Extended Data Fig. 3d, e)<sup>10,11</sup>. Suppression of 53BP1 or ligase IV strongly reduced fusion frequency (Fig. 2e) and prevented mitotic arrest (Fig. 2f, h), but did not reduce the number of deprotected telomeres (Fig. 2g and Extended Data Fig. 3f), thereby separating mitotic delay from DDR. Both telomere fusion and mitotic arrest phenotypes were suppressed by ATM inhibition<sup>12–14</sup> (Extended Data Fig. 3g–j), again indicating that telomere fusion underlies mitotic arrest. ATM inhibition did not suppress mitotic arrest induced by Taxol<sup>15</sup> (Extended Data Fig. 3k and l), confirming that the inhibitor does not perturb the SAC. Additionally, cells expressing a small hairpin RNA targeting TRF2 (shTRF2-F), which causes telomere deprotection in the absence of fusion<sup>2</sup>, did not undergo arrest (Fig. 2a, b and Extended Data Fig. 2e, f). These data are consistent with the observation that senescent cells, while harbouring a number of unfused deprotected telomeres<sup>2,16</sup> do not display mitotic arrest (Fig. 1b and Extended Data Fig. 1a).

Deletion of *TRF2* increased anaphase bridge frequency and pericentrin foci (Extended Data Fig. 4a, b) 7 days post infection, indicating multipolar mitosis when cells display telomere fusions and mitotic arrest. Accordingly, sgTRF2-2 cells exhibit unaligned metaphase chromosomes (Extended Data Fig. 4c), suggesting a chromosome congression defect. Tetraploidy did not increase as dramatically as the pericentrin foci (Extended Data Fig. 4d), ruling out tetraploidization and centrosome amplification due to anaphase bridges and cytokinesis failure, as the cause of multipolarity. We conclude that multipolarity contributes to the mitotic arrest phenotype, although it is not clear how telomere fusions drive centrosome abnormality.

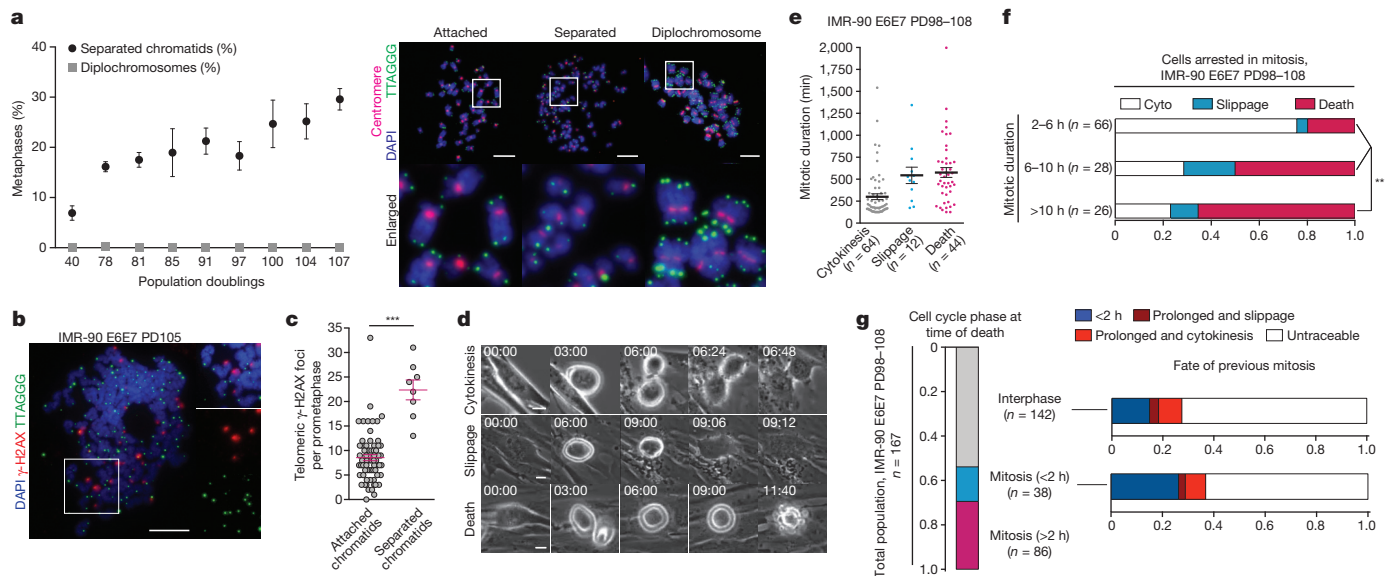
Mitotic arrest in young IMR-90 cells induces mitotic telomere deprotection<sup>3</sup>. To examine whether spontaneous mitotic arrest in pre-crisis cells also induces telomere deprotection we used premature sister separation, either cause or consequence of mitotic arrest<sup>17,18</sup>, as a marker of prolonged mitosis. Indeed, cells in pre-crisis displayed

increased premature chromatid separation (Fig. 3a). Telomeric  $\gamma$ -H2AX foci analysis (metaphase telomere-dysfunction-induced foci, meta-TIF)<sup>19,20</sup> revealed that metaphase spreads with separated chromatids displayed increased telomere deprotection (Fig. 3b, c and Extended Data Fig. 5a). Accordingly, suppression of TRF2 in young IMR90 E6E7 cells not only caused fusion and prolonged mitosis (Fig. 2a, b), but also increased premature separation (Extended Data Fig. 5b).

Cells exhibiting premature separation were mostly near-diploid, excluding the possibility that they have an increased number of telomeric ends (Extended Data Fig. 5c). We conclude that prolonged mitosis in pre-crisis is associated with a telomere DDR, and that mitotic arrest occurs in near-diploid cells. Consistently the ratio of tetraploid cells in pre-crisis cultures did not increase as prominently as the percentage of cells undergoing mitotic arrest (Extended Data Fig. 5d, e). We also rarely observed diplochromosomes, which are a consequence of two rounds of DNA replication without mitosis (endoreduplication) and thereby tetraploid (Fig. 3a).

Next we asked how prolonged mitotic arrest affects cellular fate in pre-crisis. Live cell imaging indicated three potential outcomes: cytokinesis, slippage (mitotic exit without cytokinesis), and cell death (Fig. 3d)<sup>21</sup>. Mitotic duration correlated with cellular fate, as cells that spent the least amount of time in mitosis underwent cytokinesis, and cells that resided longer in mitosis tended to slip or die (Fig. 3e). The cell death ratio increased significantly from 19% in cells arrested for 2–6 h to 50% in cells arrested for 6–10 h and to 69% in cells arrested for >10 h (Fig. 3f).

However, we noted that cells also died during interphase. Live-cell imaging revealed that 32% (86/266) of deaths occurred after prolonged mitosis (>2 hours), 14% (38/266) of cells died during mitosis lasting <2 hours and 53% (142/266) of cells died in interphase (Fig. 3g). To address whether cells that died during a short mitosis or interphase were associated with prolonged mitosis in the previous cell cycle, we traced the cells in question to the mitosis before death. Of the cells that succumb in interphase, 46% of the traceable previous mitosis was prolonged (18/39) (Fig. 3g). Of the cells that died during a short mitosis, 29% of the previous mitosis



**Figure 3 | Cell fate decision during telomere crisis.** **a**, Mean percentage  $\pm$  s.d. of metaphases with separated chromatids or diplochromosomes in IMR-90 E6E7 cells ( $n = 3$ , >100 metaphases per experiment). For one-way ANOVA separated chromatids,  $P < 0.0001$ ; diplochromosomes, not significant. Right, representative images. Scale bar, 10  $\mu$ m. **b**, **c**, Meta-TIF analysis of pre-crisis IMR-90 E6E7 cells. Scale bar, 10  $\mu$ m. Scatter plots show mean telomeric  $\gamma$ -H2AX foci  $\pm$  s.e.m. on attached chromatids ( $n = 67$ ) or separated

chromatids ( $n = 8$ ), \*\*\* $P < 0.0001$ , Mann-Whitney tests. **d**, The three distinct fates of prolonged mitosis. **e**, **f**, Mean mitotic duration of prolonged mitosis associated with the indicated fate (**e**) and ratio of each fate after indicated duration of mitotic arrest (**f**) in the pre-crisis IMR-90 E6E7 cells shown in Fig. 1b. \*\* $P < 0.005$ , Fisher's exact test. **g**, Cell cycle phase at cell death in pre-crisis IMR-90 E6E7 cells (left) and fate of the previous mitosis before death in interphase or mitosis (<2 h) (right).



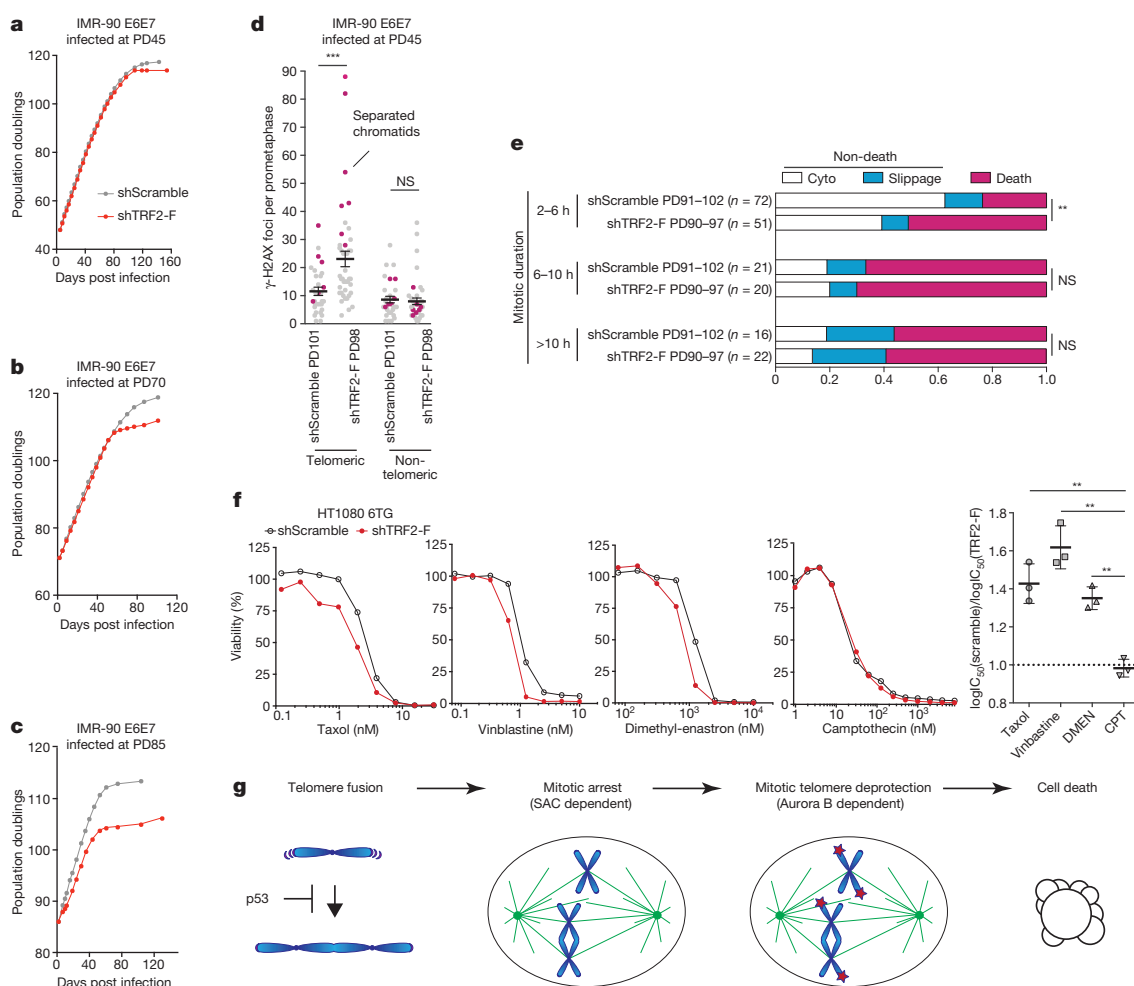
was prolonged (4/14) (Fig. 3g), indicating that cell death during either short mitosis or in interphase follows prolonged mitosis in the previous cycle.

Partial knockdown of *TRF2* exacerbates mitotic telomere deprotection in young IMR-90 E6E7 cells<sup>2</sup>. To understand whether mitotic telomere deprotection is the cause of death in pre-crisis, we tested whether a partial knockdown of *TRF2* would enhance cell death upon spontaneous mitotic arrest. IMR-90 E6E7 cells were transduced with shTRF2-F at PD45, PD70 or PD85<sup>2</sup> (Extended Data Fig. 6a). The resulting partial suppression of TRF2 did not affect cell growth dynamics in the younger populations, but slowed down growth in the population infected at PD85 (Fig. 4a–c). However, in all settings, the cells entered crisis prematurely (Fig. 4a–c).  $\gamma$ -H2AX foci analysis around PD100 revealed that shTRF2-F cells suffer from increased numbers of TIF, which was greatly enhanced on separated chromatids (Fig. 4d and Extended Data Fig. 6b, c). Accordingly, live-cell imaging revealed shTRF2-F expression increased the cell death ratio, especially after a short period of mitotic arrest (Fig. 4e). In contrast, *TRF2* overexpression (Extended Data Fig. 6d), which partially suppresses mitotic telomere deprotection<sup>3</sup>, reduced cell death after mitotic arrest, suppressed TIF and delayed crisis (Extended Data Fig. 6e–g). Neither shTRF2-F nor *TRF2* overexpression affected mitotic duration (Extended Data Fig. 6h, i). shTRF2-F also did not increase fusion frequency (Extended Data Fig. 6j), thereby attributing the increased

cell death to exaggerated loss of TRF2, not to increased fusion formation. Accordingly, loss of TRF2 in cells lacking 53BP1 significantly increased cell death and shortened the mitotic duration before cell death upon colcemid-induced mitotic arrest (Extended Data Figs 3d and 7a, b). Furthermore, inhibition of Aurora B kinase by hesperadin, which suppresses mitotic telomere deprotection<sup>3</sup> (Extended Data Fig. 7c), greatly reduced cell death upon colcemid exposure (Extended Data Fig. 7d, e), supporting a model in which amplified telomere deprotection induced by mitotic arrest triggers cell death.

This model predicts that shTRF2-F would sensitize cells to drugs that induce mitotic arrest. Therefore HT1080 6TG cells expressing shScramble or shTRF2-F were exposed to Taxol, vinblastine, dimethyl-enastron<sup>22</sup>, and as a control the topoisomerase I inhibitor camptothecin. Expression of shTRF2-F significantly sensitized HT1080 6TG cells to Taxol, vinblastine, and dimethyl-enastron, all of which induce mitotic arrest (Fig. 4f). No such effect was observed upon exposure to camptothecin (Fig. 4f), confirming that TRF2 levels affect cellular fate specifically upon mitotic arrest.

Here we show that chromosome end-to-end fusions during crisis cause spontaneous mitotic arrest, amplifying telomere deprotection, which determines cellular fate. We suggest that telomere deprotection upon spontaneous mitotic arrest is the underlying molecular signal that leads to cell death in crisis (Fig. 4g). While we cannot rule out a role of fusion breakage cycles and the resulting chromosomal



**Figure 4 | Mitotic telomere deprotection dictates cellular fate upon mitotic arrest.** **a–c**, Growth curves of indicated cells. **d**, Scatter plots show mean telomeric and non-telomeric  $\gamma$ -H2AX foci  $\pm$  s.e.m. in individual indicated cells shown in **a** ( $n > 30$ ). IF-FISH, as in Fig. 3b. Magenta, metaphases with separated chromatids. **e**, Ratio of indicated fates in pre-crisis IMR-90 E6E7 cells expressing shScramble and shTRF2-F infected at PD85, analysed as in Fig. 3f

(Fisher's exact test). **f**, Viability assay of HT1080 6TG expressing shScramble and shTRF2-F. Right, the ratio between logIC<sub>50</sub> of shScramble and shTRF2-F from three independent experiments (two-tailed *t*-test). **g**, Model of mitotic cell death pathway during crisis. \* $P < 0.05$ , \*\* $P < 0.005$ , \*\*\* $P < 0.0001$ . NS, not significant.

abnormalities as a cause of death<sup>23</sup>, several observations argue against it. Only few fusions are observed in crisis cells (Fig. 2a), which is enough to trigger prolonged mitosis (Fig. 2b). Disruption of TRF2 leads to the rapid onset of prolonged mitosis, arguing against long-term effects of fusion breakage cycles (Fig. 2a). Cells succumb to death in the first prolonged mitosis, or in the following cell cycle (Fig. 3f, g), ruling out a long-term process. Increasing damage signals without increasing fusion frequency causes more death (Fig. 4e and Extended Data Fig. 7a), arguing for signalling from deprotected telomeres as the cause for death. As most cell death during pre-crisis is associated with mitotic arrest, we propose that prolonged mitosis is the main mechanism that limits cellular lifespan upon bypass of senescence. These findings might also offer a clinical opportunity, since exacerbation of mitotic telomere deprotection sensitizes cancer cells to mitotic drugs. Mitotic arrest, however, has also been associated with tumorigenesis in checkpoint-compromised cells<sup>24</sup>. Similarly, bone marrow failure and cancer in individuals with telomeropathies are frequent, which could potentially be explained by mitotic arrest resulting from overly short telomeres<sup>25,26</sup>. Telomere-driven spontaneous mitotic arrest and the resulting mitotic telomere deprotection in the pre-crisis stage may thus function as a double-edged sword, explaining both cell death and chromosome instability upon bypass of senescence.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 October 2014; accepted 30 April 2015.

- Wright, W. E., Pereira-Smith, O. M. & Shay, J. W. Reversible cellular senescence: implications for immortalization of normal human diploid fibroblasts. *Mol. Cell. Biol.* **9**, 3088–3092 (1989).
- Cesare, A. J., Hayashi, M. T., Crabbe, L. & Karlseder, J. The telomere deprotection response is functionally distinct from the genomic DNA damage response. *Mol. Cell* **51**, 141–155 (2013).
- Hayashi, M. T., Cesare, A. J., Fitzpatrick, J. A., Lazzerini-Denchi, E. & Karlseder, J. A telomere-dependent DNA damage checkpoint induced by prolonged mitotic arrest. *Nature Struct. Mol. Biol.* **19**, 387–394 (2012).
- Le Poole, I. C. *et al.* Generation of a human melanocyte cell line by introduction of HPV16 E6 and E7 genes. *In Vitro Cell. Dev. Biol. Anim.* **33**, 42–49 (1997).
- Bodnar, A. G. *et al.* Extension of life-span by introduction of telomerase into normal human cells. *Science* **279**, 349–352 (1998).
- Meyerson, M. *et al.* hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell* **90**, 785–795 (1997).
- Santaguida, S., Tighe, A., D'Alise, A. M., Taylor, S. S. & Musacchio, A. Dissecting the role of MPS1 in chromosome biorientation and the spindle checkpoint through the small molecule inhibitor reversine. *J. Cell Biol.* **190**, 73–87 (2010).
- Hauf, S. *et al.* The small molecule hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint. *J. Cell Biol.* **161**, 281–294 (2003).
- Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
- Dimitrova, N., Chen, Y.-C. M., Spector, D. L. & de Lange, T. 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. *Nature* **456**, 524–528 (2008).
- Smogorzewska, A., Karlseder, J., Holtgreve-Grez, H., Jauch, A. & de Lange, T. DNA ligase IV-dependent NHEJ of deprotected mammalian telomeres in G1 and G2. *Curr. Biol.* **12**, 1635–1644 (2002).
- Denchi, E. L. & de Lange, T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* **448**, 1068–1071 (2007).
- Guo, X. *et al.* Dysfunctional telomeres activate an ATM-ATR-dependent DNA damage response to suppress tumorigenesis. *EMBO J.* **26**, 4709–4719 (2007).
- Hickson, I. *et al.* Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. *Cancer Res.* **64**, 9152–9159 (2004).
- Schiff, P. B. & Horwitz, S. B. Taxol stabilizes microtubules in mouse fibroblast cells. *Proc. Natl Acad. Sci. USA* **77**, 1561–1565 (1980).
- Kaul, Z., Cesare, A. J., Huschtscha, L. I., Neumann, A. A. & Reddel, R. R. Five dysfunctional telomeres predict onset of senescence in human cells. *EMBO Rep.* **13**, 52–59 (2011).
- Hoque, M. T. & Ishikawa, F. Cohesin defects lead to premature sister chromatid separation, kinetochore dysfunction, and spindle-assembly checkpoint activation. *J. Biol. Chem.* **277**, 42306–42314 (2002).
- Stevens, D., Gassmann, R., Oegema, K. & Desai, A. Uncoordinated loss of chromatid cohesion is a common outcome of extended metaphase arrest. *PLoS ONE* **6**, e22969 (2011).
- Takai, H., Smogorzewska, A. & de Lange, T. DNA damage foci at dysfunctional telomeres. *Curr. Biol.* **13**, 1549–1556 (2003).
- Cesare, A. J. *et al.* Spontaneous occurrence of telomeric DNA damage response in the absence of chromosome fusions. *Nature Struct. Mol. Biol.* **16**, 1244–1251 (2009).
- Rieder, C. L. & Maiato, H. Stuck in division or passing through: what happens when cells cannot satisfy the spindle assembly checkpoint. *Dev. Cell* **7**, 637–651 (2004).
- Gartner, M. *et al.* Development and biological evaluation of potent and specific inhibitors of mitotic Kinesin Eg5. *ChemBioChem* **6**, 1173–1177 (2005).
- Gisselsson, D. *et al.* Telomere-mediated mitotic disturbances in immortalized ovarian epithelial cells reproduce chromosomal losses and breakpoints from ovarian carcinoma. *Genes Chromosom. Cancer* **42**, 22–33 (2005).
- Dalton, W. B. & Yang, V. W. Role of prolonged mitotic checkpoint activation in the formation and treatment of cancer. *Future Oncol.* **5**, 1363–1370 (2009).
- Armanios, M. & Blackburn, E. H. The telomere syndromes. *Nature Rev. Genet.* **13**, 693–704 (2012).
- Batista, L. F. & Artandi, S. E. Understanding telomere diseases through analysis of patient-derived iPS cells. *Curr. Opin. Genet. Dev.* **23**, 526–533 (2013).

**Acknowledgements** All data are archived at the Salk Institute. We thank the Salk Institute's J. Fitzpatrick of the Waitt Advanced Biophotonics Center and members of GT3 Core, C. O'Shea, G. Wahl, F. Zhang, and D. Sabatini for support and Karlseder laboratory members for comments. M.T.H. was supported by the Human Frontier Science Program and the Japan Society for the Promotion of Science Postdoctoral Fellowships for Research Abroad. A.J.C. was supported by a NIH NRSA T32 Fellowship (5T32CA009370). T.R. was supported by the Glenn Center for Research on Aging and CIRM training grant TG2-01158. The Salk Institute Cancer Center Core Grant (P30CA014195), the NIH (R01GM087476, R01CA174942), the Donald and Darlene Shiley Chair, the Highland Street Foundation, the Fritz B. Burns Foundation, the Emerald Foundation and the Glenn Center for Research on Aging support J.K.

**Author Contributions** M.T.H. and A.J.C. designed and performed experiments, and wrote the manuscript. T.R. performed experiments, J.K. designed experiments and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.K. ([karlseder@salk.edu](mailto:karlseder@salk.edu)).

## METHODS

**Cell culture and treatment.** Human IMR-90 primary lung fibroblasts (ATCC) and their derivatives were grown in Glutamax-DMEM (Gibco) supplemented with 0.1 mM nonessential amino acids and 15% fetal bovine serum. HT1080 6TG (ATCC) cells were grown in Glutamax-DMEM supplemented with 0.1 mM nonessential amino acids and 10% bovine growth serum. All cells were grown at 7.5% CO<sub>2</sub> and 3% O<sub>2</sub>. Colcemid (Life Technologies), Taxol (A. G. Scientific), vinblastine (A. G. Scientific), dimethylnastron (A. G. Scientific), hesperadin (Selleck), reversine (Selleck) and camptothecin (Selleck) were used at the indicated concentrations. ATM inhibitor (KU-55933) (Tocris) was used at 10  $\mu$ M. FACS analysis was performed as described<sup>27</sup>. Cells were tested for mycoplasma contamination and found negative.

**Live-cell imaging.** Live-cell imaging was performed in 8-well microslide chambers (iBidi) on a Zeiss inverted fluorescent microscope with a 20 $\times$  0.8 NA air objective at 37 °C and 7.5% CO<sub>2</sub> (XLmulti S1 module, Zeiss). Images were captured with an AxioCam MRm (Zeiss) using Axio Vision software (Zeiss) typically every 6 minutes for at least 48 h. Mitotic duration was defined as movie frames from nuclear envelope breakdown or a previous frame of cell rounding to a frame of cytokinesis, slippage or mitotic cell death. Prolonged mitosis was defined as a mitosis that continues for more than 2 h. Cells that escaped from a movie screen during prolonged mitosis were included in the mitotic duration analysis but excluded from the cell fate analysis.

For shTRF2-F, image capture was started 5 days after infection. For sgEMPTY (guide RNA without target), sgTRF2-1 and sgTRF2-2, image capture was started 7 days after infection. Where indicated, population doubling is the one at their seeding, typically 1 day before the starting date of image capture. Typically more than two independent movies were analysed to confirm reproducibility except for Fig. 1b, where one movie was analysed per data point. Where indicated, Taxol, colcemid and hesperadin were added to the culture right before starting live imaging.

**Vectors and viral infections.** Target sequences of CRISPR/Cas9 are as follows<sup>28</sup>: sgTRF2-1, 5'-ACTGCATAACCCGAGCAAT-(PAM)-3'; sgTRF2-2, 5'-TGTC TGTCGCGGATTGAAGA-(PAM)-3'; sg53BP1, 5'-CAGAATCATCCTCTAG AACC-(PAM)-3'; and sgLIG4, 5'-TGGCGTCGAAACATACTAG-(PAM)-3'. Target sequences of sg53BP1 and sgLIG4 were first cloned into LentiCRISPR vector (Addgene plasmid 49535), followed by recloning of the guide RNA expression cassette (U6 promoter, target sequence and gRNA scaffold) into NheI site of LentiCas9-Blast vector (Addgene plasmid 52962). BsmBI-digested LentiCRISPR vector and LentiGuide-puro (Addgene plasmid 52963) was incubated with T4 PNK (NEB), Klenow fragment (NEB), and then T4 DNA ligase (NEB) to generate LentiCRISPR-sgEMPTY and LentiGuide-sgEMPTY, respectively. Silent mutations that pWZL-TRF2<sup>RsgRNA</sup> carries are: 5'-GCTCCTCAGAGTGATGCAAT-3' for sgTRF2-1; and 5'-TGCCTCAGCAGAATCGAGGA-3' for sgTRF2-2.

IMR-90 cells were infected with pLXSN3-HPV16E6, pLXSN3-HPV16E7, pLXSN3-HPV16E6E7, pLXSN3-p53dd, and pBabe-hTert retroviral vectors as described<sup>29</sup>, and subjected to long-term culturing in the presence of 600  $\mu$ g ml<sup>-1</sup> G418 (Mediatech, Inc.) (for pLXSN3) or 2  $\mu$ g ml<sup>-1</sup> puromycin (Mediatech, Inc.) (for pBabe). TRF2<sup>RsgRNA</sup>-expressing IMR-90 E6E7 cells were selected in the presence of 100 ng ml<sup>-1</sup> hygromycin (Mediatech, Inc.) for 6 days and subjected to LentiCRISPR infection.

Lentiviral vectors pseudotyped with VSV glycoprotein were generated by the Salk Institute Gene Transfer, Targeting and Therapeutics (GT3) Core using a modified protocol<sup>30</sup>. For shScramble (Addgene plasmid 1864), shTRF2-F, LentiCRISPR-sgEMPTY, LentiCRISPR-sgTRF2-1, LentiCRISPR-sgTRF2-2, LentiCas9-Blast, LentiCas9-Blast-sg53BP1, LentiCas9-Blast-sgLIG4, LentiGuide-sgEMPTY, and LentiGuide-sgTRF2-2, cells were plated in growth media containing 4  $\mu$ g ml<sup>-1</sup> polybrene and lentivirus and cultured for 2 days. Puromycin and blasticidin were added to the culture at 2  $\mu$ g ml<sup>-1</sup> and 10  $\mu$ g ml<sup>-1</sup>, respectively, and infected cells were selected for more than 3 days before analysis. Where

indicated, ATMi was added 4 days after infection and medium was refreshed every 1–2 days. For fusion assay, cells were harvested 7 days after infection or at indicated PD.

**Antibodies.** Primary antibodies: anti- $\gamma$ -H2AX (613402 clone 2F3, Biolegend); anti-TRF1 (Karlseder lab); anti-TRF2 (Karlseder lab); anti-53BP1 (H-300) (sc-22760, Santa Cruz); anti-LIG4 (ab80514, Abcam); anti-pericentrin (ab4448, Abcam); anti-MPM-2 (05-368MG, Millipore); anti- $\gamma$ -tubulin (T6557, Sigma-Aldrich); anti-H3S10P (D2C8) (3377, Cell Signaling); anti-GAPDH (A300-641A, Bethyl); anti-p53 (sc-126, Santa Cruz).

Secondary antibodies: HRP-linked anti-mouse or anti-rabbit (NXA931 or NA934V; GE Healthcare); Alexa488-conjugated anti-rabbit (Invitrogen); Alexa594-conjugated anti-mouse (Invitrogen).

**Western blotting.** Western blots were performed as described previously<sup>3</sup>.

**Immunofluorescence and telomere-centromere FISH on metaphase spreads.**

Cytopentrifugation, immunofluorescence and telomeric FISH were performed as described previously<sup>20</sup>. For telomere and centromere double-staining, Alexa488-conjugated telomeric PNA probe (TelC-A488, PNA Bio Inc.) and Cy3-conjugated centromeric PNA probe (CENT-Cy3, PNA Bio Inc.) were used. Percentage of telomere fusion per chromosome end was analysed as described<sup>20</sup>. For pericentrin foci analysis, cells in early mitosis (pro-, prometa- and metaphase) were selected according to chromosome shape from MPM-2-positive cells. For chromosome alignment analysis, cells in which most H3S10-P-positive chromosomes align between (and do not overlap with) two  $\gamma$ -tubulin foci were selected as metaphase. Cells were analysed 7 days after infection or at indicated PD.

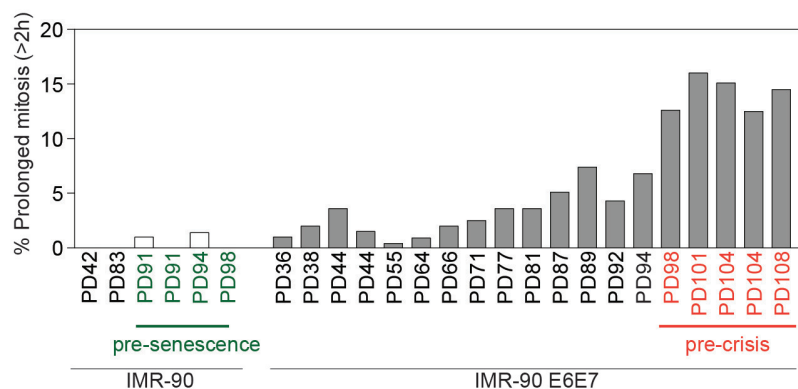
**Telomere blots.** Double-stranded telomere analysis was performed as described previously<sup>29</sup>.

**Viability assay.** Premixed WST-1 Cell Proliferation Reagent (Clontech) was used for viability assay according to manufacturer's instruction. HT1080 6TG cells infected with either shScramble or shTRF2 were seeded in 96-well plates at 6 days post infection and exposed to Taxol, vinblastine, dimethyl-enastron, and camptothecin at 7 days post infection for 48 h. Triplicate wells were analysed for each drug concentration. The results were reproduced by three independent experiments. The logIC<sub>50</sub> value was analysed by log(inhibitor) versus normalized response (variable slope) method using Prism 6 software.

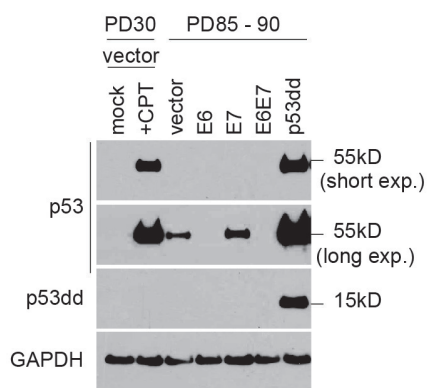
**Statistical methods.** Each figure legend shows the number of samples per experiment and number of experiments that were analysed independently. No statistical methods were used to predetermine sample size. Two-tailed unpaired *t*-tests and two-tailed Mann–Whitney tests were used to compare two data sets, where Gaussian distribution is assumed and not assumed, respectively. To detect any trends among multiple data sets in Fig. 3a and Extended Data Figs 5a and 7c, one-way ANOVA was used. Mitotic duration data sets of IMR-90 E6E7 cells in Fig. 1b were analysed with Kruskal–Wallis tests to detect trends ( $P < 0.0001$ ), in addition to Mann–Whitney tests, as described above. For statistical analysis of cellular fate after mitotic arrest, the ratio of death versus non-death (cytokinesis and slippage) was analysed with a two-tailed Fisher's exact test. For Fig. 3f, data from short (2–6 h) and middle (6–10 h) mitotic arrest were combined and compared to that of long mitotic arrest (>10 h). The null hypothesis was rejected when *P* values were less than 0.05. No randomization was performed. The investigators were not blinded to allocation during experiments and outcome assessment. All statistical analysis was performed using Prism 6 software.

27. Karlseder, J. p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* **283**, 1321–1325 (1999).
28. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
29. Karlseder, J. Senescence induced by altered telomere state, not telomere loss. *Science* **295**, 2446–2449 (2002).
30. Tiscornia, G., Singer, O. & Verma, I. M. Production and purification of lentiviral vectors. *Nature Protocols* **1**, 241–245 (2006).

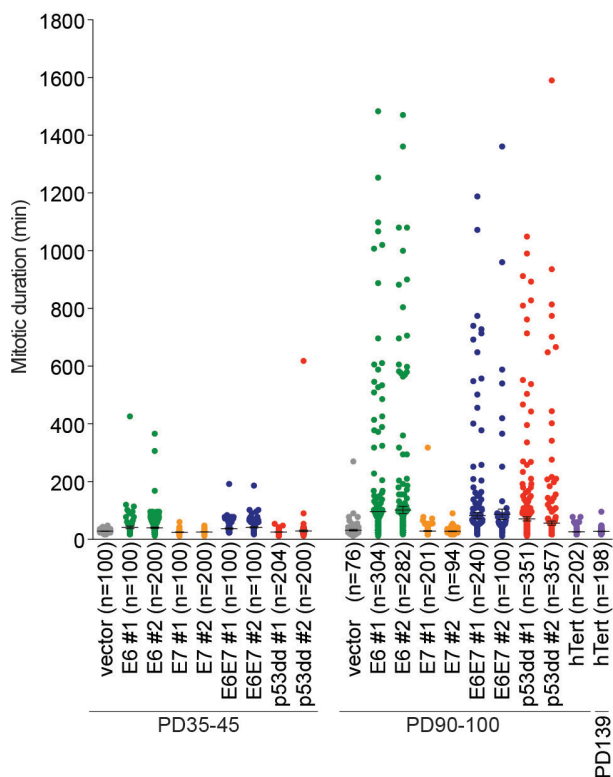
a



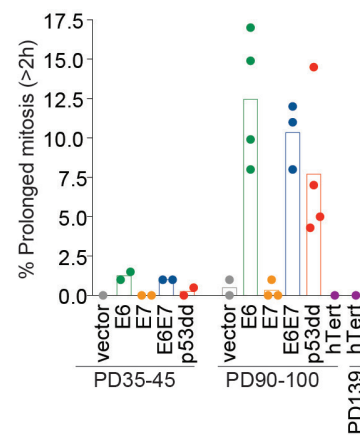
b



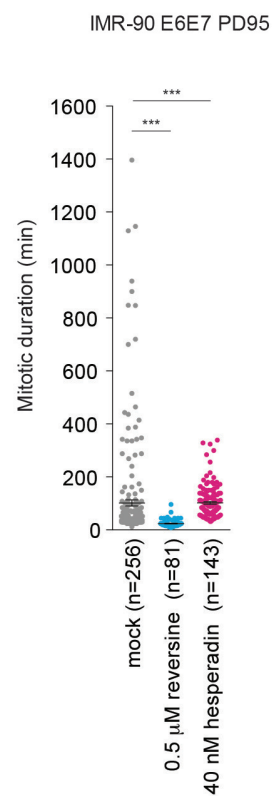
c



d



e

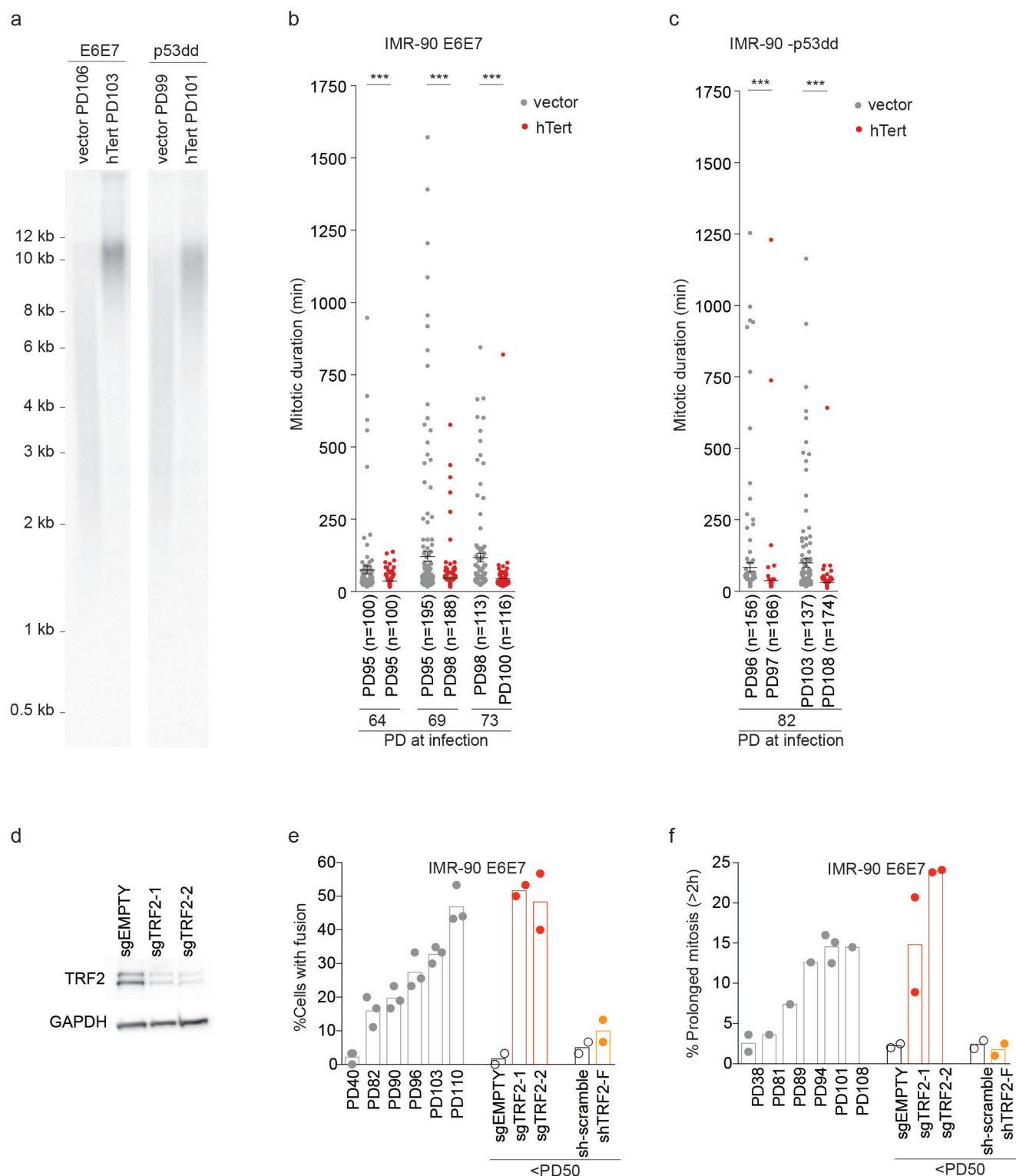




**Extended Data Figure 1 | Mitotic duration in pre-crisis cells is elongated.**

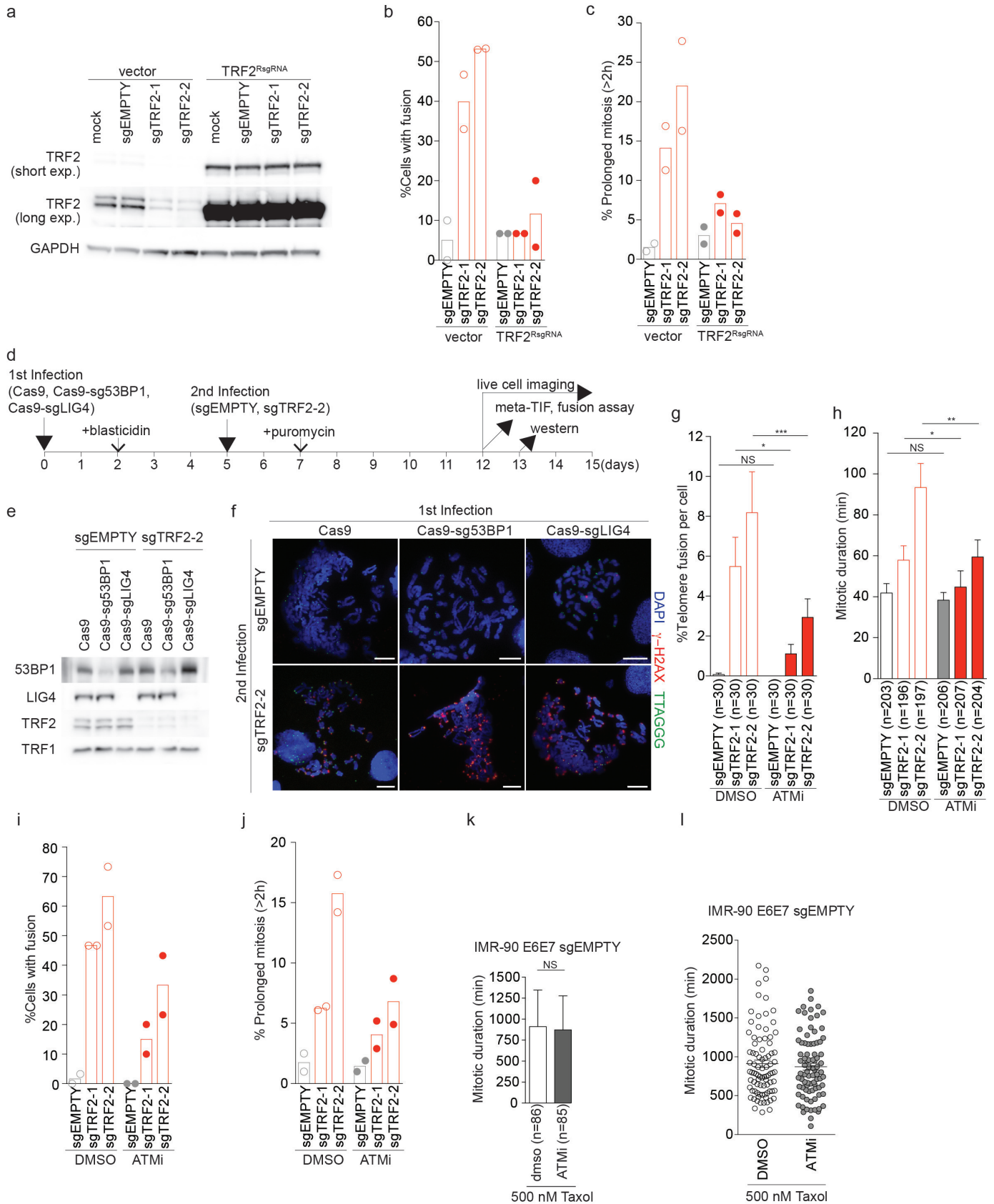
**a**, Percentage of cells that spend more than 2 h in mitosis (prolonged mitosis), shown in Fig. 1b. **b**, Effect of indicated oncogenes on p53 expression. **c**, Scatter plots show mean mitotic duration  $\pm$  s.e.m. of individual IMR-90 derivative cells analysed in Fig. 1d. **d**, Scatter plots with bars show mean percentage of

prolonged mitosis analysed in Fig. 1d (1–4 independent experiments). The dots represent independent experiments and the bars the mean. **e**, Scatter plots show mean ( $\pm$  s.e.m.) mitotic duration of IMR-90 E6E7 PD95 cells exposed to reversine and hesperadin (\*\* $P < 0.0001$ , Mann–Whitney test). The result was reproduced in two independent experiments.



**Extended Data Figure 2 | Telomere elongation abrogates extended mitotic duration.** **a**, Telomere elongation by hTert in IMR-90 E6E7 and p53dd cells shown in Fig. 1e, f. IMR-90 E6E7 and p53dd cells were infected at PD73 and PD82, respectively, and analysed at the indicated PD. **b**, **c**, Scatter plots show mean mitotic duration  $\pm$  s.e.m. of individual IMR-90 E6E7 (**b**) and p53dd (**c**)

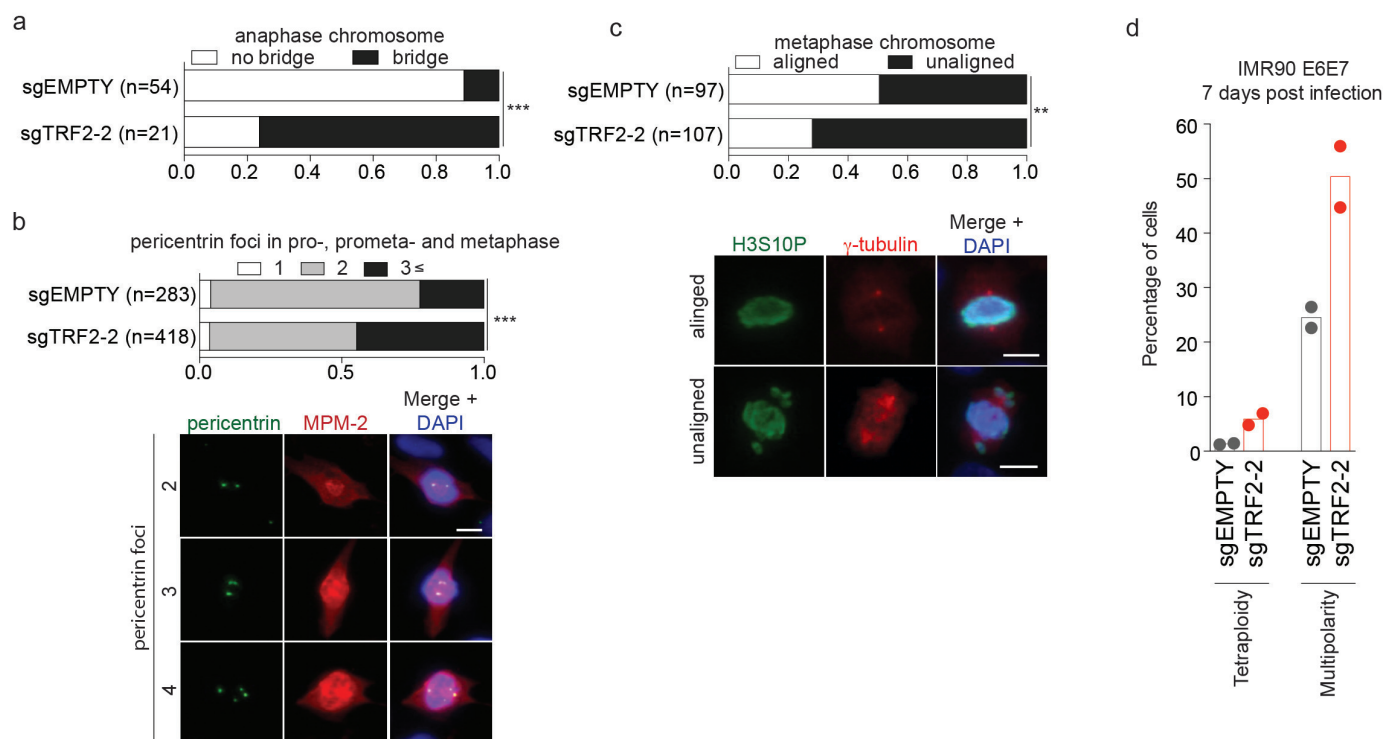
cells expressing hTert at indicated PD shown in Fig. 1e, f. **d**, Effect of sgTRF2 on TRF2 expression 7 days after infection. **e**, **f**, Scatter plots with bars show mean percentage of cells with telomeric fusion (**e**) and prolonged mitosis (**f**) in IMR-90 E6E7 derivatives shown in Fig. 2a, b (1–3 independent experiments). \*\*\* $P < 0.0001$ . Mann–Whitney test.



**Extended Data Figure 3 | Extended mitotic duration depends on telomere fusion.** **a**, Effect of sgTRF2 on TRF2 expression in cells expressing sgTRF2-resistant TRF2 (TRF2<sup>RsgRNA</sup>) 9 days after infection with CRISPR/Cas9. **b, c**, Scatter plots with bars show mean percentage of cells with telomeric fusion (**b**) and prolonged mitosis (**c**) in IMR-90 E6E7 expressing sgTRF2 in the presence of TRF2<sup>RsgRNA</sup> shown in Fig. 2c, d (two independent experiments). **d**, Schematic of 53BP1 or ligase IV suppression experiment in the presence of sgEMPTY or sgTRF2-2. **e**, Western analysis of IMR90 E6E7 cells expressing Cas9, Cas9-sg53BP1 or Cas9-sgLig4 in the background of sgEMPTY or sgTRF2-2. **f**, Representative meta-TIF images of cells suppressed for 53BP1 or LIG4 in the presence of sgEMPTY or sgTRF2-2 as described in **d**.

**g, h**, Percentage of telomeric fusion (**g**) and mitotic duration (**h**) in IMR-90 E6E7 expressing sgEMPTY and sgTRF2 in the presence of DMSO or ATM inhibitor (mean  $\pm$  s.e.m.). **i, j**, Scatter plots with bars showing mean percentage of cells with telomeric fusion (**i**) and prolonged mitosis (**j**) in IMR-90 E6E7 expressing sgTRF2 in the presence of ATM inhibitor shown in **g, h** (two independent experiments). **k**, Mean mitotic duration  $\pm$  s.e.m. of IMR-90 E6E7 sgEMPTY cells exposed to 500 nM Taxol in the presence of DMSO or ATMi. **l**, Scatter plots show mean mitotic duration of individual cells shown in **j**. NS, not significant; \* $P < 0.05$ , \*\* $P < 0.005$ , \*\*\* $P < 0.0001$ , Mann-Whitney tests. Results were reproduced in at least two independent experiments.

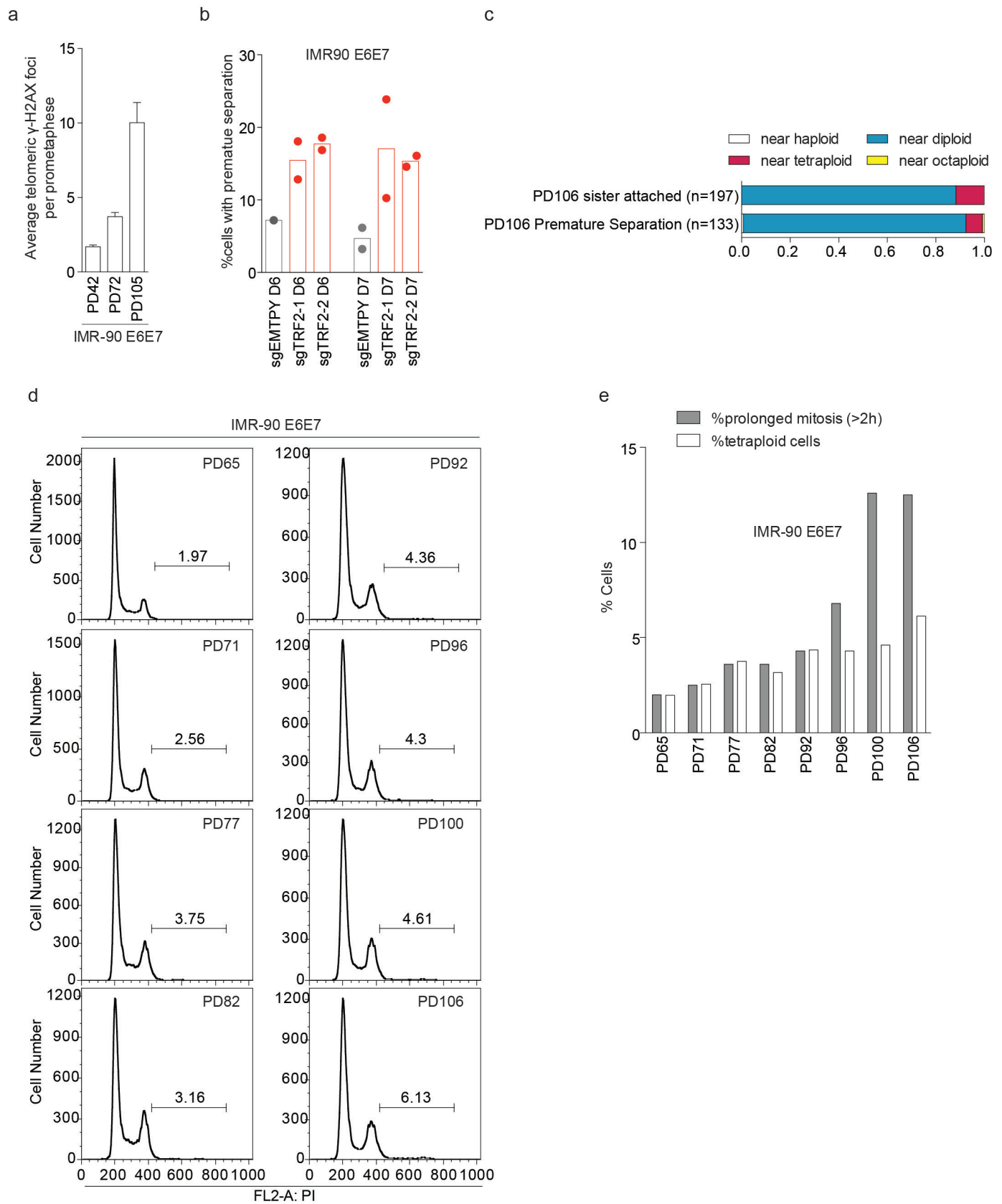




#### Extended Data Figure 4 | Telomere fusions lead to multipolar mitosis.

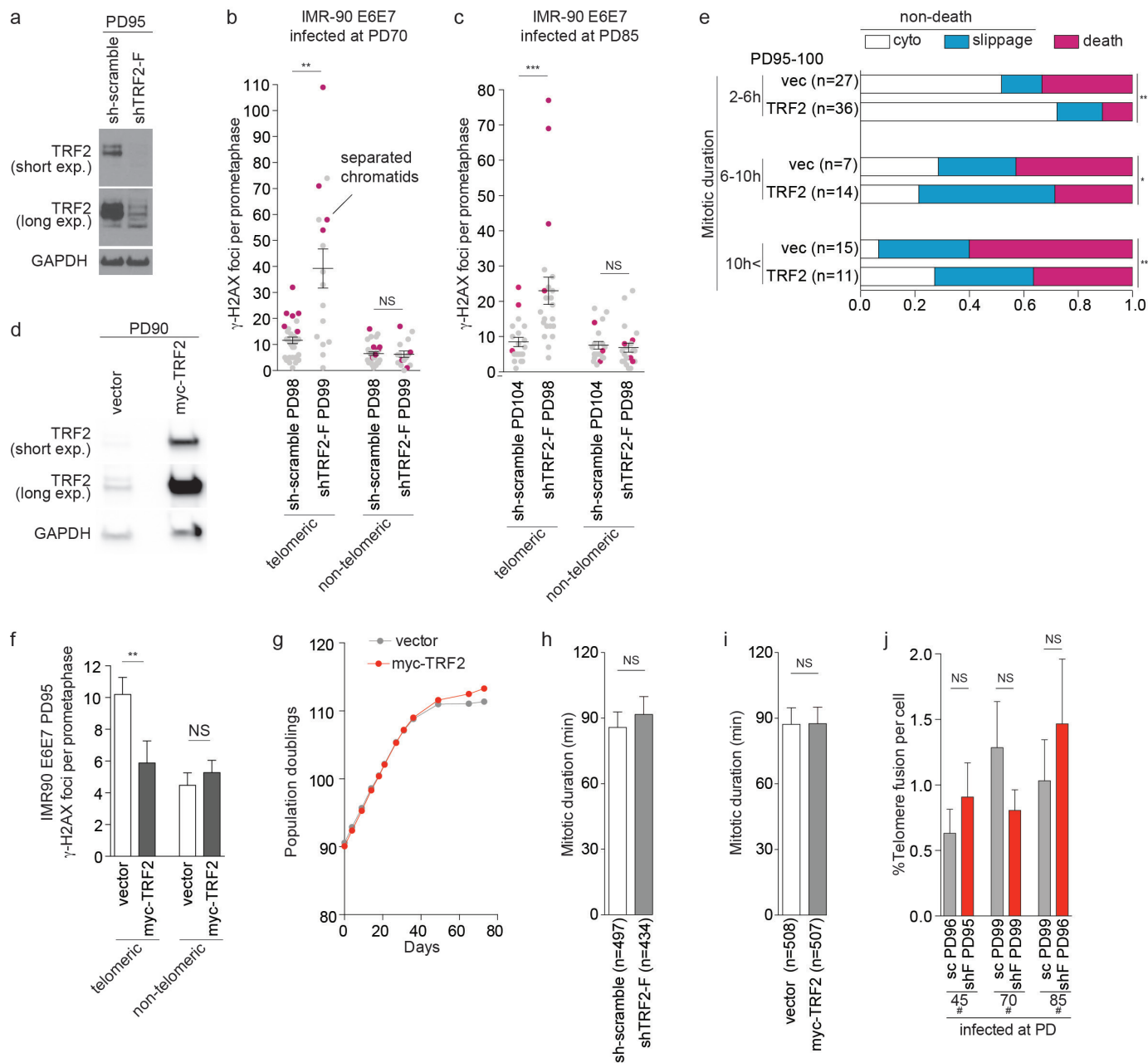
**a–c**, Ratio of anaphase chromosome with or without anaphase bridge (**a**), pericentrin foci in pro-, prometa- and metaphase (**b**) and metaphase chromosome with or without unaligned chromosome (**c**) in IMR-90 E6E7 expressing sgEMPTY and sgTRF2-2 7 days after infection (Fisher's exact test, for pericentrin foci, 1 and 2 foci versus  $\geq 3$  foci). Representative images from

sgEMPTY cells are shown below (**b**, **c**). Results were reproduced in two independent experiments. **d**, Scatter plots with bars show mean percentage of cells that possess tetraploidy (FACS analysis) and multipolarity ( $\geq 3$  pericentrin foci as in **j**) in IMR-90 E6E7 cells expressing sgEMPTY and sgTRF2-2 7 days after infection (two independent experiments). \*\* $P < 0.005$ , \*\*\* $P < 0.0001$ . Fisher's exact test. Scale bars, 10  $\mu\text{m}$ .



**Extended Data Figure 5 | Mitotic arrest occurs in near-diploid cells.** **a**, Bars show mean of three independent experiments of average number ( $\pm$  s.e.m.) of telomeric  $\gamma$ -H2AX foci in IMR-90 E6E7 at PD42, 72 and 105 as analysed in Fig. 4d ( $n = 25$  per experiment,  $P = 0.0008$ , one-way ANOVA). **b**, Bars show the percentage of young IMR90 E6E7 cells with separated sister chromatids expressing sgEMPTY or sgTRF2-1/2 6 or 7 days post infection with CRISPR/Cas9. Results from two independent experiments except for sgEMPTY 6 days

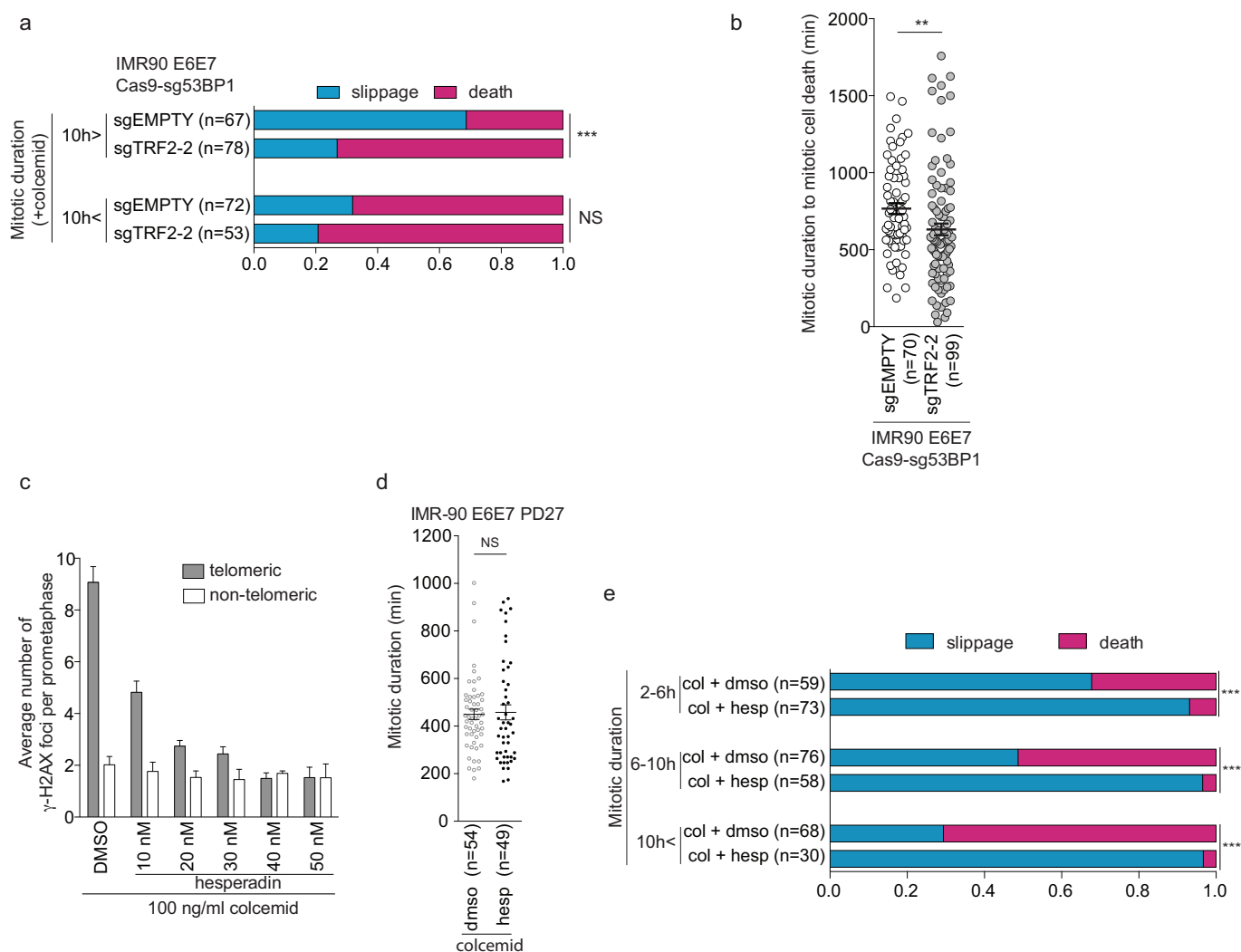
post infection (at least 31 metaphases per experiment) are shown. **c**, Ratio of ploidy of mitotic IMR-90 E6E7 cells at PD106. Cells with attached and separated sister chromatids are plotted separately. **d**, FACS analysis of cells shown in Fig. 1b. **e**, Bars show percentage of prolonged mitosis and tetraploid cells ( $4N$ ) in IMR-90 E6E7 cells at indicated PD. Percentage of prolonged mitosis are the same data sets as Extended Data Fig. 1a.



**Extended Data Figure 6 | TRF2 suppression and overexpression affects cell fate.** **a**, Effect of shTRF2-F on TRF2 expression 7 days after infection. **b, c**, Telomeric and non-telomeric  $\gamma$ -H2AX foci in individual pre-crisis IMR-90 E6E7 cells expressing shScramble and shTRF2-F shown in Fig. 4b, c. Data in **b** and **c** were analysed as in Fig. 4d ( $n > 16$ , mean  $\pm$  s.e.m.). Metaphases with separated chromatids are shown in magenta. **d**, Effect of Myc-TRF2 on TRF2 expression 7 days after infection. **e**, Ratio of mitotic cell fate in pre-crisis IMR-90 E6E7-expressing TRF2 cells, analysed as in Fig. 3f (Fisher's exact test, death versus non-death). **f**, Meta-TIF analysis of pre-crisis IMR-90 E6E7 cells expressing a control vector or Myc-TRF2. Bars show mean ( $\pm$  s.e.m.)

telomeric and non-telomeric  $\gamma$ -H2AX foci ( $n = 25$ , Mann-Whitney tests). **g**, Growth curve of IMR-90 E6E7 cells expressing a control vector or infected with Myc-TRF2-F at PD90. **h, i**, Mean ( $\pm$  s.e.m.) mitotic duration of IMR-90 E6E7 cells expressing shScramble and shTRF2-F (**h**) and Myc-TRF2 (**i**) analysed in Fig. 4e and Extended Data Fig. 6e, respectively (Mann-Whitney test). **j**, Mean ( $\pm$  s.e.m.) percentage of telomere fusion of pre-crisis IMR-90 E6E7 expressing shScramble and shTRF2-F analysed in Fig. 4a-c (Mann-Whitney test). \* $P < 0.05$ , \*\* $P < 0.005$ , \*\*\* $P < 0.0001$ . NS, not significant.





**Extended Data Figure 7 | Amplified telomere deprotection during mitotic arrest in crisis causes cell death.** **a**, Ratio of mitotic slippage and cell death in IMR-90 E6E7 Cas9-sg53BP1 cells expressing sgEMPTY and sgTRF2-2 in the presence of colcemid, analysed as in Fig. 3f (Fisher's exact test, death versus slippage). **b**, Scatter plots show mean mitotic duration  $\pm$  s.e.m. before cell death of individual IMR-90 E6E7 Cas9-sg53BP1 cells expressing sgEMPTY and sgTRF2-2 in the presence of colcemid (Mann-Whitney test). **c**, Bars show mean ( $\pm$  s.d.) of three independent experiments of average telomeric and non-telomeric  $\gamma$ -H2AX foci in IMR-90 E6E7 cells at PD45 exposed to colcemid in

the presence of DMSO or hesperadin at indicated concentrations for 24 h analysed as in Fig. 4d (50 metaphase per experiment). For one-way ANOVA telomeric foci,  $P < 0.0001$ ; non-telomeric foci, not significant. **d**, Scatter plots show mean ( $\pm$  s.e.m.) mitotic duration of IMR-90 E6E7 cells at PD27 exposed to 100 ng ml<sup>-1</sup> colcemid in the presence of DMSO or 40 nM hesperadin (Mann-Whitney test). **e**, Ratio of mitotic cell fate in IMR-90 E6E7 cells around PD45 exposed to 100 ng ml<sup>-1</sup> colcemid in the presence of DMSO or 40 ng ml<sup>-1</sup> hesperadin (Fisher's exact test, death versus slippage). \*\* $P < 0.005$ , \*\*\* $P < 0.0001$ . NS, not significant.

# New cofactor supports $\alpha,\beta$ -unsaturated acid decarboxylation via 1,3-dipolar cycloaddition

Karl A. P. Payne<sup>1</sup>, Mark D. White<sup>1</sup>, Karl Fisher<sup>1</sup>, Basile Khara<sup>1</sup>, Samuel S. Bailey<sup>1</sup>, David Parker<sup>2</sup>, Nicholas J. W. Rattray<sup>1</sup>, Drupad K. Trivedi<sup>1</sup>, Royston Goodacre<sup>1</sup>, Rebecca Beveridge<sup>1</sup>, Perdita Barran<sup>1</sup>, Stephen E. J. Rigby<sup>1</sup>, Nigel S. Scrutton<sup>1</sup>, Sam Hay<sup>1</sup> & David Leys<sup>1</sup>

The bacterial *ubiD* and *ubiX* or the homologous fungal *fdc1* and *pad1* genes have been implicated in the non-oxidative reversible decarboxylation of aromatic substrates, and play a pivotal role in bacterial ubiquinone (also known as coenzyme Q) biosynthesis<sup>1–3</sup> or microbial biodegradation of aromatic compounds<sup>4–6</sup>, respectively. Despite biochemical studies on individual gene products, the composition and cofactor requirement of the enzyme responsible for *in vivo* decarboxylase activity remained unclear<sup>7–9</sup>. Here we show that Fdc1 is solely responsible for the reversible decarboxylase activity, and that it requires a new type of cofactor: a prenylated flavin synthesized by the associated UbiX/Pad1<sup>10</sup>. Atomic resolution crystal structures reveal that two distinct isomers of the oxidized cofactor can be observed, an isoalloxazine N5-iminium adduct and a N5 secondary ketimine species with markedly altered ring structure, both having azomethine ylide character. Substrate binding positions the dipolarophile enoic acid group directly above the azomethine ylide group. The structure of a covalent inhibitor–cofactor adduct suggests that 1,3-dipolar cycloaddition chemistry supports reversible decarboxylation in these enzymes. Although 1,3-dipolar cycloaddition is commonly used in organic chemistry<sup>11,12</sup>, we propose that this presents the first example, to our knowledge, of an enzymatic 1,3-dipolar cycloaddition reaction. Our model for Fdc1/UbiD catalysis offers new routes in alkene hydrocarbon production or aryl (de)carboxylation.

Decarboxylation is among one of the most common reactions in nature, despite the fact it is inherently difficult to achieve at ambient conditions. This is due to the high energy of the carbanion intermediate that is formed concomitant with carbon dioxide formation. To overcome this challenge, the majority of decarboxylases<sup>13</sup> make use of cofactors, including organic cofactors such as flavins, pyridoxal phosphate (PLP), thiamine pyrophosphate (TPP) or metal ions such as Mg<sup>2+</sup>, Fe<sup>2+</sup> or Mn<sup>2+</sup>. Decarboxylation is also frequently coupled to substrate oxidation. Comparatively few decarboxylases have been shown to require no cofactor and in these selected cases catalysis involves non-oxidative decarboxylation of those substrates for which the corresponding carbanion species can be stabilized using simple acid-base chemistry. Examples include orotidine monophosphate decarboxylase<sup>14</sup> and arylmalonate decarboxylase<sup>15</sup>.

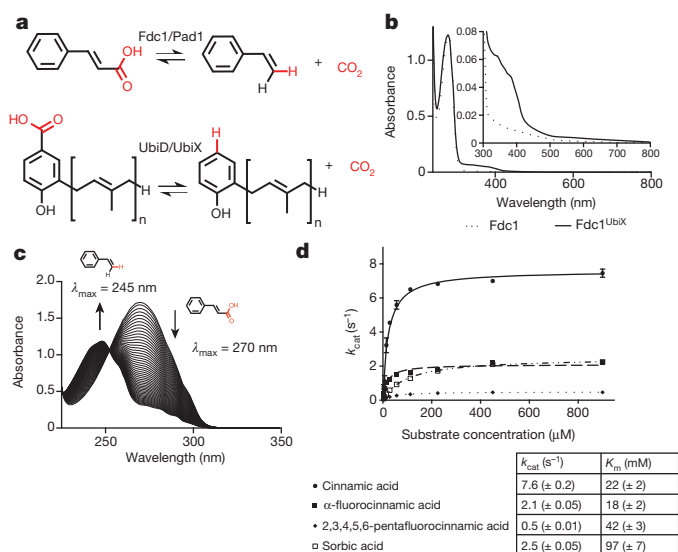
The *ubiX* and *ubiD* or the related *pad1* and *fdc1* genes have been shown to be responsible for non-oxidative reversible decarboxylation of aromatic substrates<sup>1–6</sup> (Fig. 1a). These genes are widely distributed in archaeal, bacterial and fungal genomes<sup>16</sup>, and *ubiX* and *ubiD* feature in the prokaryotic ubiquinone biosynthetic pathway<sup>1–3</sup>. Genetic studies have led to the suggestion that both genes encode (redundant) decarboxylases<sup>17</sup>. However, although UbiX/Pad1 proteins are distantly related to flavin-containing cysteine decarboxylases and have been shown to bind FMN, no *in vitro* decarboxylase activity has been detected<sup>7,18</sup>. Furthermore, most biochemical studies of UbiD/Fdc1 have also failed to detect *in vitro* activity<sup>8,19</sup>. Although recent UbiD/Fdc1

structures reveal distant structural homology to a family of NADH–FMN oxidoreductases, only metal ion binding rather than flavin binding has been reported<sup>8</sup>. Hence, the composition and putative cofactor requirement of the enzyme responsible for the observed *in vivo* decarboxylase activity remained unclear.

We co-expressed the *Aspergillus niger fdc1* gene in *E. coli* with either the associated *A. niger pad1* or the homologous *ubiX* from *E. coli*. Although no direct interaction could be detected between the purified Fdc1 and either UbiX or Pad1 proteins, clear differences could be observed between single expressed Fdc1 and the corresponding *ubiX* or *pad1* co-expressed Fdc1 protein (denoted Fdc1<sup>UbiX</sup>). Purified Fdc1<sup>UbiX</sup> has distinct spectral properties (Fig. 1b) and catalyses reversible decarboxylation of a wide range of aromatic carboxylic acids *in vitro* (Fig. 1c, d and Extended Data Fig. 1). Similar observations were made for *Saccharomyces cerevisiae* Fdc1<sup>UbiX</sup>, although in this case the corresponding Fdc1 singly expressed protein (but not Fdc1<sup>UbiX</sup>) weakly binds FMN (Extended Data Fig. 2). During the review stage for this publication, similar observations for the *S. cerevisiae* Fdc1 were reported elsewhere, although the Fdc1 cofactor was not identified<sup>20</sup>.

Crystal structures of *A. niger*, *Candida dubliniensis* and *S. cerevisiae* Fdc1<sup>UbiX</sup> reveal a heavily modified FMN cofactor is bound by these enzymes (Fig. 2 and Extended Data Fig. 3). The modified FMN phosphate is bound in complex with metal ions, similar to the distantly related FMN-binding protein from *Methanobacterium thermoautotrophicum*<sup>21</sup>. The electron density suggests both Mn<sup>2+</sup> and K<sup>+</sup> ions are present, and Mn<sup>2+</sup> binding can be detected using EPR (Extended Data Fig. 4). Early studies on *E. coli* UbiD revealed activation by both Mn<sup>2+</sup> and an unidentified cofactor<sup>3</sup>. The electron density for the modified FMN reveals extensive modification has occurred at both N5 and C6 positions, effectively adding a fourth ring to the isoalloxazine moiety (Fig. 2b). The atomic resolution obtained for *A. niger* Fdc1<sup>UbiX</sup> allows identification of the cofactor structure as an isopentenyl-adduct to the flavin N5–C6 or prenylated FMN (prFMN) (Fig. 2c). The branched nature of the isopentenyl adduct and the position of the covalent linkages with the flavin suggest this modification is achieved through prenylation. In the accompanying paper<sup>10</sup>, we show that UbiX indeed acts as a flavin prenylase enzyme that can support Fdc1 activation *in vitro*. Activation is dependent on the presence of oxygen, suggesting the reduced prFMN UbiX product is oxidized to the corresponding flavo-N5 iminium adduct (prFMN<sup>iminium</sup>; Fig. 2d). The presence of the oxidized prFMN was confirmed by high-resolution mass spectrometry of both the isolated cofactor and native mass spectrometry of the Fdc1<sup>UbiX</sup> protein (Extended Data Fig. 5). Furthermore, complete inactivation of Fdc1<sup>UbiX</sup> was achieved using the mild reductant sodium cyanoborohydride. Upon reoxidation of the sodium cyanoborohydride inactivated enzyme, a radical species was formed with spectral and EPR properties similar to that of the prFMN• radical species detected during non-physiological oxidation of the UbiX–prFMN complex<sup>10</sup> (Fig. 2e and Extended Data Fig. 4).

<sup>1</sup>Centre for Synthetic Biology of Fine and Speciality Chemicals, Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, UK. <sup>2</sup>Innovation/Biobrain, Shell International Exploration and Production, Westhollow Technology Center, 3333 Highway 6 South, Houston, Texas 77082-3101, USA.

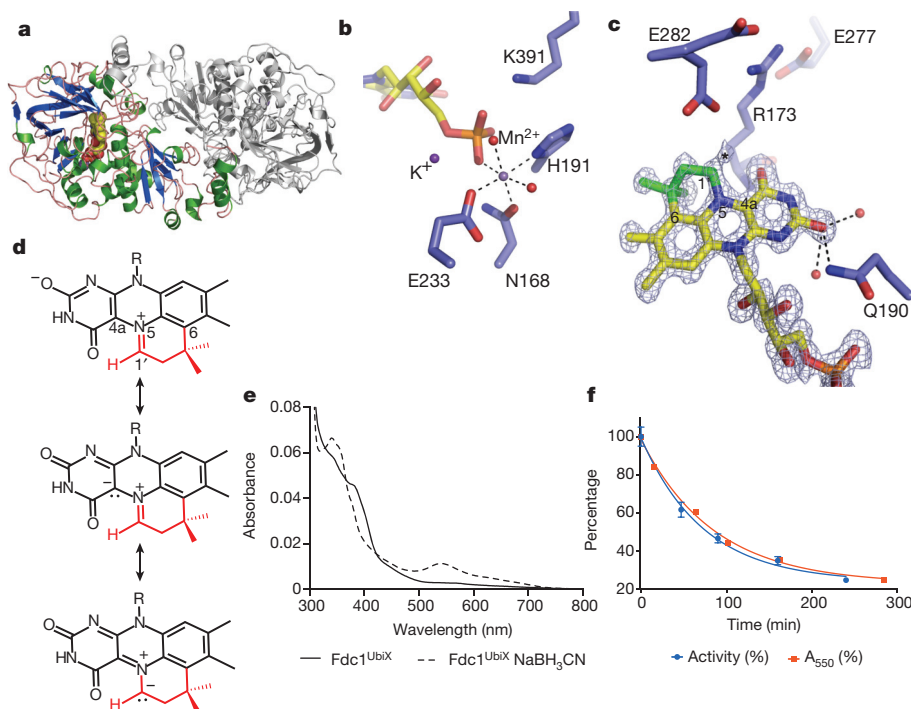


**Figure 1 | Fdc1 solution data.** **a**, Schematic overview of the reaction catalysed by Fdc1/Pad1 and the related UbiD/UbiX proteins. **b**, UV-visible spectra obtained for heterologous expressed *A. niger* Fdc1, with and without co-expression of a *ubiX* gene. **c**, UV-visible observation of the enzymatic conversion of cinnamic acid to styrene via decarboxylation by Fdc1<sup>UbiX</sup>. The initial spectrum of cinnamic acid shows a  $\lambda_{\text{max}}$  of 270 nm. Over time, successive spectra show reduction of the 270 nm peak and appearance of a peak at 245 nm corresponding to styrene formation. **d**, Steady-state kinetic parameters obtained for Fdc1<sup>UbiX</sup> for sorbic acid and a variety of cinnamic acid-type compounds (error bars are s.e.m.,  $n = 3$ ).

Furthermore, the amount of prFMN<sup>•</sup> radical formed is proportional to the levels of enzyme activity obtained before cyanoborohydride inactivation (Fig. 2f), suggesting that enzyme inactivation occurs by modification (probably hydrolysis) of the active prFMN species.

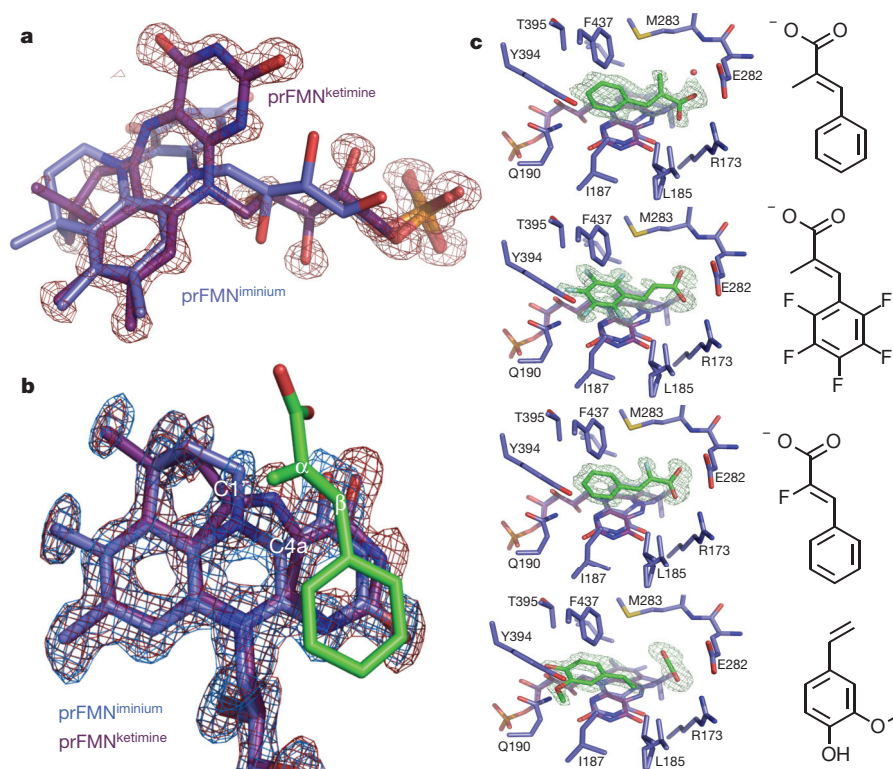
Marked heterogeneity of the prFMN<sup>ox</sup> cofactor is observed in the *A. niger* Fdc1<sup>UbiX</sup> wild-type crystal structures, with older crystal samples revealing an increased population of a hydroxylated prFMN<sup>iminium</sup>. The latter species could also be observed by high-resolution mass spectrometry (Extended Data Fig. 5a). Surprisingly, a distinct prFMN<sup>ox</sup> isomer (prFMN<sup>ketimine</sup>) can be observed for most *A. niger* Fdc1 crystals, although it is not detected in the *C. dubliniensis* or *S. cerevisiae* crystals (Fig. 3a). This form contains an altered isoalloxazine ring structure, a likely consequence of isomerization of the extended ring system during oxidative maturation of the prFMN cofactor. We propose this isomerization step occurs during the stepwise oxidation of the reduced prFMN UbiX product, via a radical mechanism (Extended Data Fig. 6). The isomerization leads to insertion of the prenyl C1' carbon into the flavin isoalloxazine ring, creating a central seven-membered ring and converting the N5-iminium into a secondary ketimine. The conformation derived from density functional theory (DFT) calculations for the proposed prFMN<sup>ketimine</sup> species is in close agreement with that observed in the crystal structure. The calculations indicate this species is 30–40 kJ mol<sup>-1</sup> lower in gas phase free energy than the prFMN<sup>iminium</sup> species and can exist in two butterfly bent-like conformations of similar energy (Extended Data Fig. 7).

To obtain insights into the catalytic mechanism, *A. niger* Fdc1<sup>UbiX</sup> crystals were soaked with a range of *trans*-cinnamic-acid-related compounds. The corresponding enzyme–substrate complexes clearly



**Figure 2 | Fdc1<sup>UbiX</sup> crystal structures.** **a**, Crystal structure of the *A. niger* Fdc1<sup>UbiX</sup>. The monomer present in the asymmetric unit is depicted in cartoon format (helices in green, sheets in blue) while the symmetry related monomer forming the Fdc1<sup>UbiX</sup> dimer is shown in grey. **b**, Detailed view of Mn<sup>2+</sup> binding site, linking the modified FMN phosphate group to the protein. **c**, Detailed view of the modified FMN cofactor. Omit electron density map corresponding to the bound cofactor contoured at 5 $\sigma$ . A fourth ring (non-isoalloxazine derived atoms are shown in green) can clearly be observed. An asterisk indicates the position of additional weak electron density at the C1 position, which can be accounted for by partial hydrolysis. Key residues involved in polar interactions with the isoalloxazine derived

cofactor moiety are shown as sticks. **d**, Chemical structure of the modified FMN bound by *A. niger* Fdc1<sup>UbiX</sup> as derived from atomic resolution density and high-resolution mass spectrometry (Extended Data Fig. 5). Atoms derived from dimethylallyl-monophosphate are shown in red<sup>10</sup>. The azomethine ylide resonance form is shown below. **e**, Oxidation of cyanoborohydride inactivated Fdc1<sup>UbiX</sup> leads to formation of a purple species, similar in UV-visible and EPR spectral properties (Extended Data Fig. 4) to that observed for the oxidized UbiX–prFMN complex<sup>10</sup>. **f**, Following purification, Fdc1<sup>UbiX</sup> activity gradually decreases when incubated on ice, and formation of the purple prFMN radical following NaBH<sub>3</sub>CN treatment follows a similar trend (error bars are s.d.,  $n = 3$ ).



**Figure 3 | Fdc1<sup>UbiX</sup> cofactor structure and ligand complexes.** **a**, Omit electron density map corresponding to a distinct isomer of the prFMN contoured at 5 $\sigma$ . An expansion of the central ring of the isoalloxazine system can clearly be observed, with the distinct butterfly bent conformation accompanied by altered conformation of the ribityl moiety. **b**, Detailed view of prFMN<sup>ox</sup> isoforms in the Fdc1<sup>UbiX</sup>–substrate complexes. The omit electron density maps corresponding respectively to the prFMN<sup>iminium</sup> (in blue) and

prFMN<sup>ketimine</sup> (in red) species each contoured at 5 $\sigma$  are shown for the  $\alpha$ -methyl-cinnamic acid complex. **c**, A series of Fdc1<sup>UbiX</sup> substrate/product complexes. Selected active site residues of Fdc1<sup>UbiX</sup> are shown in atom colour sticks, with the omit electron density contoured at 5 $\sigma$  corresponding to  $\alpha$ -methyl-cinnamic acid, pentafluorocinnamic acid,  $\alpha$ -fluorocinnamic acid and 4-vinyl-guaiacol, respectively.

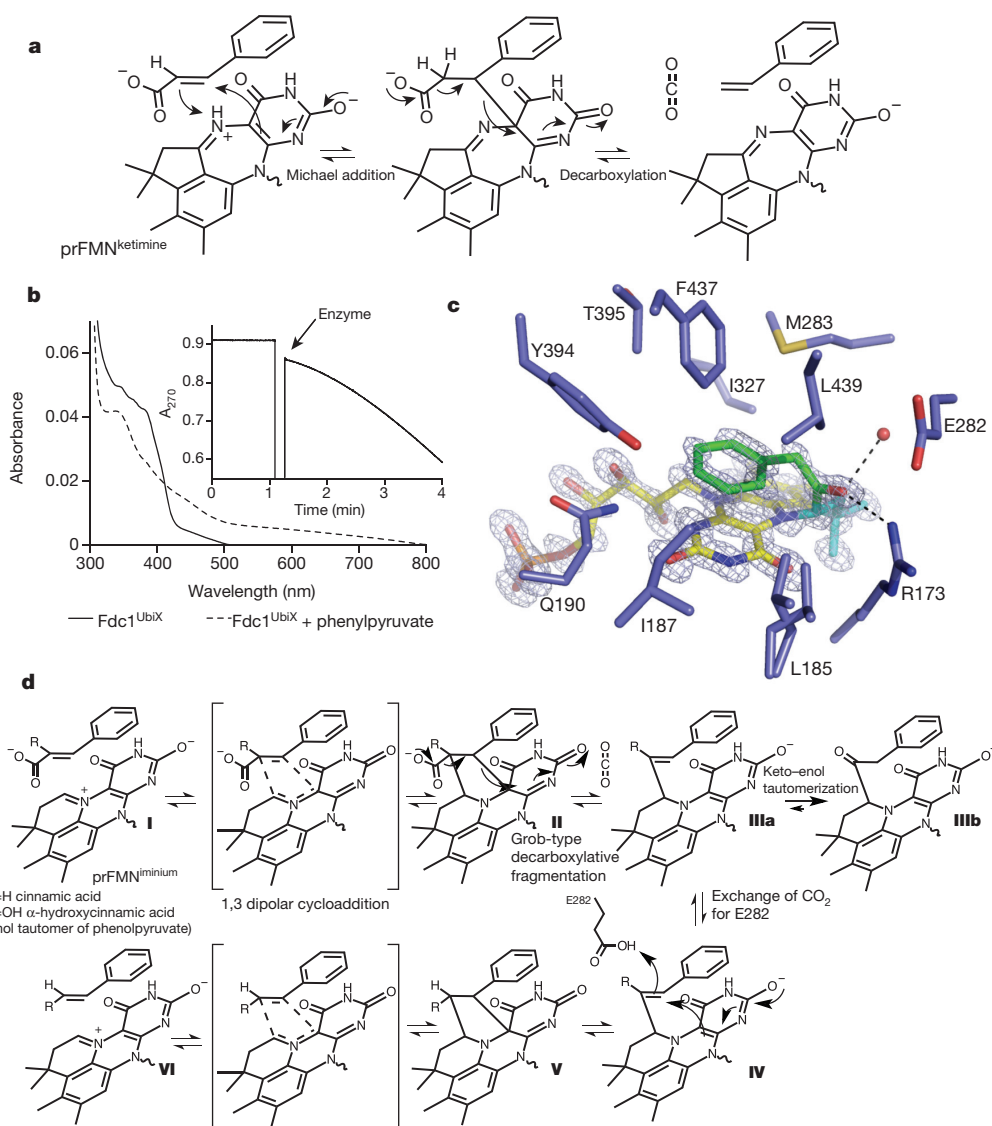
reveal substrates are stacked directly above the prFMN<sup>ox</sup> ring system (Fig. 3b, c). Electron density for both the iminium and ketimine species can be detected in the atomic resolution density maps, the latter adopting a more planar conformation of the extended ring system in order to accommodate substrate binding (Fig. 3b, c and Extended Data Fig. 7). The substrate enoic acid double bond is positioned directly above the C4a of both the prFMN<sup>ox</sup> species. Although the aryl-group of all substrates tested occupies a similar position directly above the prFMN N1 atom, electron density corresponding to the carboxylate moiety is distinct for each substrate (Fig. 3c). Although clear electron density for the pentafluorocinnamic acid carboxylate is lacking, both  $\alpha$ -methylcinnamic acid and  $\alpha$ -fluorocinnamic acid carboxylate groups can be observed. Both occupy a position vacated by reorientation of the conserved E282, and establish polar contacts with R173 and the amide nitrogen of M283. In the case of  $\alpha$ -methylcinnamic acid, the carboxylate is within hydrogen bonding distance of the N5 secondary ketimine group of the prFMN<sup>ketimine</sup> form. A complex with the decarboxylation product 4-vinyl guaiacol (a distinct flavouring agent obtained by yeast-mediated decarboxylation of ferulic acid) reveals a similar position for the aryl-moiety (Fig. 3c). The vinyl group is placed directly above the prFMN<sup>ox</sup> C4a, in close proximity to electron density resembling a bound CO<sub>2</sub> molecule.

At first glance, the position of the substrate  $\alpha,\beta$ -unsaturated carbonyl directly above the prFMN<sup>ox</sup> C4a suggests the possibility of transient formation of a prFMN<sup>ox</sup> C4a substrate  $\beta$ -carbon bond reminiscent of Michael addition-like chemistry and other flavin-catalysed reactions<sup>22</sup>. This mechanism has similarities to the amidohydrolase-type decarboxylases, where nucleophilic attack by a metal-bound hydroxide ion is postulated to lead to transient C=C bond cleavage of an enoic acid moiety leading to decarboxylation<sup>13,23</sup>. In the case of

Fdc1, C=C bond cleavage should proceed with concomitant protonation of the  $\alpha$ -carbon to allow decarboxylation. Our structural data suggests the N5 secondary ketimine of prFMN<sup>ketimine</sup> could act as a probable acid-base catalyst, providing a rationale for both FMN modification and rearrangement (Fig. 4a and Extended Data Fig. 7b). However, although our solution mass spectrometry and chemical inhibition data cannot distinguish between the iminium and ketimine cofactor isomers, solution data presented in Fig. 2e, f suggests prFMN<sup>iminium</sup> is the catalytically relevant species.

To determine which prFMN<sup>ox</sup> isomer supports catalysis by Fdc1<sup>UbiX</sup> and provide further insight into the likely mechanism, we sought to determine the structure of a covalent substrate–cofactor adduct. When *A. niger* Fdc1<sup>UbiX</sup> protein is incubated with phenylpyruvate, for which the corresponding enol-tautomer  $\alpha$ -hydroxycinnamic acid (present as a minor population in solution) closely resembles the cinnamic acid substrate, the UV-visible spectrum is altered (Fig. 4b). Similar observations are made when incubating with phenylacetaldehyde. Furthermore, incubation with phenylpyruvate or phenylacetaldehyde leads to reversible enzyme inhibition, with a gradual increase of enzyme activity upon removal of excess phenylpyruvate (Fig. 4b). The corresponding crystal structure of *A. niger* Fdc1<sup>UbiX</sup> incubated with phenylpyruvate reveals a covalent adduct formed between the prenyl-C1' of the prFMN<sup>iminium</sup> cofactor and a substrate-derived entity (Fig. 4c). Atomic resolution electron density reveals the latter clearly corresponds to a phenylacetaldehyde adduct, a species that can be formed following decarboxylation of  $\alpha$ -hydroxycinnamic acid and tautomerisation of the corresponding  $\alpha$ -hydroxystyrene prFMN<sup>iminium</sup> adduct (Fig. 4d, species **IIIb**). The reversible inhibition observed can be accounted for by the fact the corresponding enol–keto tautomer equilibrium is in favour of the off-pathway phenylacetaldehyde adduct.





**Figure 4 | Insights into the Fdc1<sup>UbiX</sup> mechanism.** **a**, Proposed mechanism for Fdc1<sup>UbiX</sup> catalysis using the prFMN ketimine form. **b**, Incubation of Fdc1<sup>UbiX</sup> with phenylpyruvate leads to changes in the UV-visible spectrum. The inset shows conversion of cinnamic acid to styrene (monitored at 270 nm) after addition of phenylpyruvate treated Fdc1<sup>UbiX</sup>. **c**, Active site of Fdc1<sup>UbiX</sup> in

complex with a phenylpyruvate derived adduct. The omit electron density map (contoured at 7σ, in blue) corresponding to the prFMN<sup>iminium</sup>–phenylpyruvate adduct is shown. Atoms derived from phenylpyruvate are depicted in green. **d**, Proposed mechanism for Fdc1<sup>UbiX</sup> catalysis and phenylacetaldehyde adduct formation of the prFMN<sup>iminium</sup> form.

This observation further confirms that prFMN<sup>iminium</sup> is the catalytically relevant species, and the unexpected presence of a C1'–substrate α-carbon bond suggests a distinct mechanism from that of the amidohydrolase-type decarboxylases occurs (Fig. 4d). Fluorination of the substrate β-carbon or the aromatic group has only a modest effect on catalysis (Fig. 1d). Furthermore, simple enoic acids also act as substrates, as prolonged incubation of Fdc1<sup>UbiX</sup> with these substrates does lead to formation of the corresponding terminal alkenes (α-olefins) at modest levels (Extended Data Fig. 8). Both observations argue against formation of transient ionic species.

The neutral form of the prFMN<sup>iminium</sup> has distinct azomethine ylide character (a well-known 1,3-dipole) that is positioned directly adjacent to the substrate α,β-unsaturated carbonyl (a dipolarophile) by the enzyme. Thus, we propose a 1,3-dipolar cycloaddition between the substrate and the prFMN<sup>iminium</sup> azomethine ylide form leads to a transient covalent substrate–prFMN<sup>iminium</sup> pyrrolidine adduct, establishing covalent bonds with both C1' and C4a (species II). Ample precedent is available for the concerted 1,3-dipolar cycloaddition between azomethine ylides and α,β-unsaturated carbonyls in organic chemistry<sup>11,12</sup>. Whether the proposed Fdc1-mediated cycloaddition

occurs through a single pericyclic transition state cannot be established at this stage. Fragmentative decarboxylation (Grob-type fragmentation<sup>24</sup>) of the pyrrolidine adduct II can occur coupled to breaking the β-carbon–prFMN C4a bond. This leaves a single bond connecting the substrate α-carbon with the C1' of prFMN<sup>iminium</sup> (species IIIa), as observed in the phenylpyruvate adduct IIIb. We propose protonation by E282 concomitant with formation of a second pyrrolidine adduct (species V) leads to product formation (species VI) via a *retro* 1,3-dipolar cycloaddition. Together with R172 and E277, E282 forms a network of polar interactions that is conserved throughout the UbiD/Fdc1 family. The R173A, E277Q and E282Q variants of Fdc1<sup>UbiX</sup> are all inactive, and display altered UV-visible spectrum properties, suggesting these mutations also affect cofactor maturation (Extended Data Fig. 9).

The modification of the flavin isoalloxazine ring through prenylation at N5 and C6, followed by oxidation strikingly alters and expands the flavin chemical repertoire<sup>22</sup>. Oxidation of the N5-prenyl adduct bond leads either to the corresponding iminium adduct or, when coupled to ring isomerization, a secondary ketimine adduct. Both species have azomethine ylide character and distinct catalytic

potential. In the Fdc1/UbiD enzyme family, the prFMN<sup>iminium</sup> form supports reversible decarboxylation of a wide range of (aromatic) substrates. Although the [3+2] reaction between azomethine ylide dipoles and alkene dipolarophiles has been extensively used in organic chemistry<sup>11,12</sup>, the mechanism proposed here would present the first example, to our knowledge, of a biological [3+2] reaction. Combined with the recent description of both natural<sup>25</sup> and artificial<sup>26</sup> bona fide [4+2] cycloaddition catalysing enzymes, this hints at more widespread use of pericyclic reaction chemistry in nature. As Fdc1/UbiD enzymes have evolved from an NADH-FMN oxidoreductase module<sup>8,21</sup>, the distinct possibility exists that other unrelated prFMN-dependent enzymes might have developed from distinct flavin-binding modules. Such distinct prFMN-dependent enzymes could make use of different aspects of the prFMN chemistry, as occurs with other organic cofactors<sup>27</sup>. Artificial flavoenzymes containing N5-alkylated flavins have been created that are capable of H<sub>2</sub>O<sub>2</sub>-driven enantioselective sulfoxidations<sup>28</sup>, using the fact that N5-alkylated flavins can form remarkably stable 4 $\alpha$ -peroxyflavins and are powerful oxidising catalysts<sup>29</sup>. It is possible similar enzymes already exist in nature.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 8 December 2014; accepted 13 May 2015.**

**Published online 17 June 2015.**

- Aussel, L. *et al.* Biosynthesis and physiology of coenzyme Q in bacteria. *Biochim. Biophys. Acta* **1837**, 1004–1011 (2014).
- Gulmezian, M., Hyman, K. R., Marbois, B. N., Clarke, C. F. & Javor, G. T. The role of UbiX in *Escherichia coli* coenzyme Q biosynthesis. *Arch. Biochem. Biophys.* **467**, 144–153 (2007).
- Leppik, R. A., Young, I. G. & Gibson, F. Membrane-associated reactions in ubiquinone biosynthesis in *Escherichia coli*. 3-octoprenyl-4-hydroxybenzoate carboxylase. *Biochim. Biophys. Acta* **436**, 800–810 (1976).
- Erb, T. J. Carboxylases in natural and synthetic microbial pathways. *Appl. Environ. Microbiol.* **77**, 8466–8477 (2011).
- Boll, M., Loeffler, C., Morris, B. E. L. & Kung, J. W. Anaerobic degradation of homocyclic aromatic compounds via arylcarboxyl-coenzyme A esters: organisms, strategies and key enzymes. *Environ. Microbiol.* **16**, 612–627 (2014).
- Mukai, N., Masaki, K., Fujii, T., Kawamukai, M. & Iefuji, H. PAD1 and FDC1 are essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* **109**, 564–569 (2010).
- Rangarajan, E. S. *et al.* Crystal structure of a dodecameric FMN-dependent UbiX-like decarboxylase (Pad1) from *Escherichia coli* O157:H7. *Protein Sci.* **13**, 3006–3016 (2004).
- Jacewicz, A., Izumi, A., Brunner, K., Schnell, R. & Schneider, G. Structural insights into the UbiD protein family from the crystal structure of PA0254 from *Pseudomonas aeruginosa*. *PLoS ONE* **8**, e63161 (2013).
- Stratford, M. *et al.* Mapping the structural requirements of inducers and substrates for decarboxylation of weak acid preservatives by the food spoilage mould *Aspergillus niger*. *Int. J. Food Microbiol.* **157**, 375–383 (2012).
- White, M. D. *et al.* UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis. *Nature*. <http://dx.doi.org/10.1038/nature14559> (2015).
- Pellissier, H. Asymmetric 1,3-dipolar cycloadditions. *Tetrahedron* **63**, 3235–3285 (2007).
- Ess, D. H. & Houk, K. N. Theory of 1,3-dipolar cycloadditions: Distortion/Interaction and frontier molecular orbital models. *J. Am. Chem. Soc.* **130**, 10187–10198 (2008).
- Li, T., Huo, L., Pulley, C. & Liu, A. Decarboxylation mechanisms in biological system. *Bioorg. Chem.* **43**, 2–14 (2012).
- Desai, B. J. *et al.* Investigating the role of a backbone to substrate hydrogen bond in OMP decarboxylase using a site-specific amide to ester substitution. *Proc. Natl Acad. Sci. USA* **111**, 15066–15071 (2014).
- Okrasa, K. *et al.* Structure-guided directed evolution of alkenyl and arylmalonate decarboxylases. *Angew. Chem.* **48**, 7691–7694 (2009).
- Lupa, B., Lyon, D., Gibbs, M. D., Reeves, R. A. & Wiegel, J. Distribution of genes encoding the microbial non-oxidative reversible hydroxyaryl acid decarboxylases/phenol carboxylases. *Genomics* **86**, 342–351 (2005).
- Zhang, H. & Javor, G. T. Regulation of the isofunctional genes ubiD and ubiX of the ubiquinone biosynthetic pathway of *Escherichia coli*. *FEMS Microbiol. Lett.* **223**, 67–72 (2003).
- Kopec, J., Schnell, R. & Schneider, G. Structure of PA4019, a putative aromatic acid decarboxylase from *Pseudomonas aeruginosa*. *Acta Crystallogr. F* **67**, 1184–1188 (2011).
- Plumridge, A. *et al.* The decarboxylation of the weak-acid preservative, sorbic acids, is encoded by linked genes in *Aspergillus* spp. *Fungal Genet. Biol.* **47**, 683–692 (2010).
- Lin, F., Ferguson, K. L., Boyer, D. R., Lin, X. N. & Marsh, E. N. Isofunctional enzymes Pad1 and UbiX catalyse formation of a novel cofactor required by ferulic acid decarboxylase and 4-hydroxy-3-polyphenylbenzoic acid decarboxylase. *ACS Chem. Biol.* **10**, 1137–1144 (2015).
- Christendat, D. *et al.* Structural proteomics of an archaeon. *Nature Struct. Biol.* **7**, 903–909 (2000).
- Walsh, C. T. & Wenciewicz, T. A. Flavoenzymes: versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.* **30**, 175–200 (2013).
- Xu, S. *et al.* Crystal structures of isoorotate decarboxylases reveal a novel catalytic mechanism of 5-carboxyl-uracil decarboxylation and shed light on the search for DNA decarboxylase. *Cell Res.* **23**, 1296–1309 (2013).
- Prantz, K. & Mulzer, J. Synthetic applications of the carbonyl generating Grob fragmentation. *Chem. Rev.* **110**, 3741–3766 (2010).
- Kim, H. J., Ruszczycky, M. W., Choi, S.-H., Lie, Y.-N. & Liu, H.-W. Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature* **473**, 109–112 (2011).
- Preiswerk, N. *et al.* Impact of scaffold rigidity on the design and evolution of an artificial Diels-Alderase. *Proc. Natl Acad. Sci. USA* **111**, 8013–8018 (2014).
- Richter, M. Functional diversity of organic molecule enzyme cofactors. *Nat. Prod. Rep.* **30**, 1324–1345 (2013).
- de Gonzalo, G., Smit, C., Jin, J., Minnaard, A. J. & Fraaije, M. W. Turning a riboflavin-binding protein into a self-sufficient monooxygenase by cofactor redesign. *Chem. Commun.* **47**, 11050–11052 (2011).
- Imada, Y., Iida, H., Kitagawa, T. & Naota, T. Aerobic reduction of olefins by *in situ* generation of diimide with synthetic flavin catalysts. *Chemistry* **17**, 5908–5920 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** The main part of this work was supported by BBSRC grants (BB/K017802/1 with Shell and BB/M017702/1). Early studies were supported by EU grant FP-7 256808 to D.L. and N.S.S. S.H. is a BBSRC David Phillips research fellow. N.S.S. is an EPSRC Established Career Fellow and Royal Society Wolfson Award holder. We thank Diamond Light Source for access to MX beamlines (proposal number MX8997), which helped to contribute to the results presented here. We thank D. Procter (University of Manchester) for discussions. The authors acknowledge the assistance given by IT Services and the use of the Computational Shared Facility at The University of Manchester.

**Author Contributions** K.A.P.P. carried out molecular biology, biophysical and structural biology studies of *A. niger* Fdc1. B.K. carried out molecular biology experiments underpinning biophysical and structural biology studies of *S. cerevisiae* Fdc1 performed by M.D.W. K.F. and S.E.J.R. performed and analysed EPR experiments. S.H. performed DFT calculations. N.J.W.R., D.K.T. and R.G. undertook liquid chromatography–mass spectrometry of extracts and interpreted the data on substrate–product species. R.B. and P.B. performed native mass spectrometry. S.S.B. solved the *C. dubliniensis* Fdc1 structure. All authors discussed the results with N.S.S. and D.P. and all participated in writing the manuscript. D.L. initiated and directed this research.

**Author Information** Coordinates and structure factors have been deposited in the Protein Data Bank under accession numbers 4ZA4, 4ZA5, 4ZA7, 4ZA8, 4ZAB, 4ZA9, 4ZAA, 4ZAC and 4ZAD. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.L. ([david.leys@manchester.ac.uk](mailto:david.leys@manchester.ac.uk)).

## METHODS

**Cloning.** The *S. cerevisiae*, *C. dubliniensis* and *A. niger* *fdc1* genes were codon optimized and synthesized (Genscript). The *A. niger* *fdc1* gene was cloned into the NdeI and XhoI sites of pET30a and the *S. cerevisiae* and *C. dubliniensis* genes were cloned into the NdeI and XhoI sites of pET21b. *E. coli* *ubiX* was cloned into the NdeI and XhoI sites of pET21b and pET30a. *A. niger* *fdc1* pET30a was transformed into *E. coli* BL21(DE3) with and without *ubiX* pET21b and *S. cerevisiae* and *C. dubliniensis* *fdc1* pET21b were transformed into *E. coli* BL21(DE3) with and without *ubiX* pET30a.

**Mutagenesis.** Mutagenesis primers were designed using the QuikChange Primer Design Program (<http://www.genomics.agilent.com/primerDesignProgram.jsp>). PCR was performed using Phusion polymerase (NEB). Template was removed by DpnI (NEB) digestion and the PCR product transformed into *E. coli* NEB5 $\alpha$ . Once the presence of the desired mutation was confirmed by DNA sequencing, the plasmid was co-transformed with the appropriate *ubiX* construct into *E. coli* BL21(DE3).

**Protein expression and purification.** Protein was expressed in BL21(DE3) grown at 37 °C at 180 r.p.m. in LB broth supplemented with 50  $\mu\text{g ml}^{-1}$  kanamycin and/or 50  $\mu\text{g ml}^{-1}$  ampicillin. At mid-log phase cells were induced with 0.25 mM IPTG and supplemented with 1 mM  $\text{MnCl}_2$ , grown overnight at 15 °C at 180 r.p.m. and then collected by centrifugation (4 °C, 7,000g for 10 min). Cell pellets were resuspended in buffer A (200 mM NaCl, 1 mM  $\text{MnCl}_2$ , 50 mM Tris pH 7.5) supplemented with DNase, RNase, lysozyme (Sigma) and Complete EDTA-free protease inhibitor cocktail (Roche). Cells were lysed using a French press at 20,000 p.s.i. and the lysate clarified by centrifugation at 125,000g for 90 min. The supernatant was applied to a Ni-NTA agarose column (Qiagen). The column washed with 3 column volumes of buffer A supplemented with 10 mM imidazole and protein eluted in 1-ml fractions with buffer A supplemented with 250 mM imidazole. Samples were subjected to SDS-PAGE analysis and fractions found to contain the purified protein were pooled. Imidazole was removed using a 10-DG desalting column (Bio-Rad) equilibrated with 100 mM NaCl, 1 mM  $\text{MnCl}_2$ , 25 mM Tris pH 7.5. Protein was aliquoted and flash frozen until required.

**UV-visible spectroscopy/protein quantification.** UV-visible absorbance spectra were recorded with a Cary UV-vis spectrophotometer. The protein concentration of *A. niger* Fdc1 and *S. cerevisiae* Fdc1 was estimated using  $\epsilon_{280} = 68,870 \text{ M}^{-1} \text{ cm}^{-1}$  and  $\epsilon_{280} = 63,830 \text{ M}^{-1} \text{ cm}^{-1}$  (calculated from the primary amino acid sequence using the ProtParam program on the ExPASy proteomics server), respectively.

**UV-visible decarboxylation assays.** For substrates with two or more double bonds conjugated to the acid group initial rates were determined by following consumption of substrate by UV-visible spectroscopy using a Cary 50 Bio spectrophotometer (Varian). Assays were performed against various concentrations of substrate in 350  $\mu\text{l}$  50 mM KCl, 50 mM NaPi pH 6 in a 1 mm path length cuvette at 25 °C. The rate of 2,3,4,5,6-penta-fluorocinnamic acid consumption was monitored at 260 nm,  $\alpha$ -fluorocinnamic acid at 265 nm, cinnamic acid at 270 nm and sorbic acid at 260 nm. Extinction coefficients for each substrate were calculated from a range of standards and found to be  $\epsilon_{260} = 20,000 \text{ M}^{-1} \text{ cm}^{-1}$  for 2,3,4,5,6-penta-fluorocinnamic acid;  $\epsilon_{265} = 20,000 \text{ M}^{-1} \text{ cm}^{-1}$  for  $\alpha$ -fluorocinnamic acid;  $\epsilon_{270} = 20,000 \text{ M}^{-1} \text{ cm}^{-1}$  for cinnamic acid; and  $\epsilon_{260} = 24,000 \text{ M}^{-1} \text{ cm}^{-1}$  for sorbic acid, the latter two values being in good agreement with the literature<sup>30,31</sup>.

***A. niger* Fdc1 phenylpyruvate inhibition.** Sodium phenylpyruvate powder was added to the protein until there was no further change in the  $\sim A_{380}$  spectrum. Any precipitate was removed by centrifugation at 16,100g and excess phenylpyruvate in solution removed using a 10-DG desalting column (BioRad).

***A. niger* Fdc1 inactivation timecourse.** Enzyme was thawed and incubated on ice. The enzyme was assayed against cinnamic acid at various time points and the level of activity expressed as a percentage of that of freshly thawed protein (0 min). At various time points samples were also treated with  $\text{NaBH}_3\text{CN}$  powder, centrifuged at 16,100g for 15 min and UV-visible spectra recorded.

**GC headspace decarboxylation assays.** A total of 500  $\mu\text{l}$  10 mM substrate (2-hexenoic acid, 2-heptenoic acid, 2-octenoic or 2-nonenic acid), 50 mM KCl, 50 mM NaPi pH 6 in a 2 ml GC vial was incubated with or without enzyme at 25 °C with continuous shaking. 0.5 ml head space volume was manually drawn off, and injected into the GC with a syringe-lock needle. A Varian 3800 GC equipped with a DB-WAX column (30 m  $\times$  0.32 mm  $\times$  0.25  $\mu\text{m}$  film thickness, JW Scientific) was used to detect and quantify 1-alkenes produced during the enzyme reactions. The column temperature was programmed as follows: 40 °C hold for 2 min, to 150 °C at 20 °C  $\text{min}^{-1}$ . The injector temperature was 250 °C (10:1 split), and the FID temperature was set at 250 °C. The carrier gas was helium at a flow rate of 1  $\text{ml min}^{-1}$ . Product identification was achieved by comparison with pure 1-alkene standards.

**Carboxylation HPLC assays.** Assays containing 10 mM styrene, 100 mM NaPi pH 6, 3 M  $\text{NaHCO}_3$  (final pH 7.5), were incubated with and without enzyme at 25 °C overnight. The sample was centrifuged at 16,100g to remove precipitate and

100  $\mu\text{l}$  added to 900  $\mu\text{l}$  50% v/v  $\text{H}_2\text{O}$ /acetonitrile. Sample analysis was performed using an Agilent 1100 Series HPLC equipped with a UV detector. The stationary phase was a Kinetex 5  $\mu\text{m}$  C18 100A column, 250  $\times$  4.6 mm. The mobile phase was acetonitrile/water (50/50) with 0.1% TFA at a flow rate of 1  $\text{ml min}^{-1}$ , and detection was performed at a wavelength of 254 nm.

**EPR.**  $\text{NaBH}_3\text{CN}$  powder was added to Fdc1<sup>ubiX</sup> (215  $\mu\text{M}$ ). The samples were incubated on ice and UV-visible spectra recorded until the purple colour ( $A_{550}$ ) reached a maximum. The samples were centrifuged to remove any remaining powder or effervescence, 300  $\mu\text{l}$  of sample was transferred to an EPR tube and frozen in liquid nitrogen. EPR spectra were obtained using a Bruker ELEXSYS E500 EPR spectrometer equipped with a Super High Q (ER 4118-SHQ) resonator. Temperature control was achieved using an Oxford Instruments ESR900 cryostat connected to an ITC 503 temperature controller from the same manufacturer.

**Protein crystallization and structure determination.** Crystallization was performed by vapour diffusion. An initial screening of 0.3  $\mu\text{l}$  14  $\text{mg ml}^{-1}$  *A. niger* Fdc1 in 100 mM NaCl, 25 mM Tris pH 7.5 and 0.3  $\mu\text{l}$  reservoir solution at 4 °C resulted in a number of hits including PACT condition F4 (molecular dimensions). Seed stock produced from these crystals was used in an optimization screen based around 0.2 M potassium thiocyanate, Bis-Tris propane 6.5, 20% w/v PEG 3350 mixing 0.05  $\mu\text{l}$  seed stock, 0.25  $\mu\text{l}$  protein solution and 0.3  $\mu\text{l}$  reservoir solution at 4 °C. For *S. cerevisiae* Fdc1, crystals were obtained using 0.3  $\mu\text{l}$  2  $\text{mg ml}^{-1}$  in 20 mM potassium phosphate, 200 mM NaCl pH 7.5 and 0.3  $\mu\text{l}$  of 0.2 M calcium acetate, 20% PEG 3000 at 4 °C. Crystals for *C. dubliniensis* were obtained at 4 °C using 0.3  $\mu\text{l}$  18  $\text{mg ml}^{-1}$  *C. dubliniensis* Fdc1 in 100 mM NaCl, 1 mM  $\text{MnCl}_2$ , 25 mM Tris pH 7.5 and 0.3  $\mu\text{l}$  of 0.2 M potassium thiocyanate, sodium acetate pH 5.5 and 15% w/v PEG 4000.

All crystals were cryoprotected in mother liquor supplemented with 10% PEG 200 and flash cooled in liquid nitrogen. Diffraction data was collected at Diamond beamlines at 100K and processed using the CCP4 suite<sup>32</sup>. Data was reduced and scaled using XDS<sup>33</sup>. Following molecular replacement using 4IP2 as a starting model, the structures of wild-type *A. niger* and *S. cerevisiae* Fdc1 and complexes were refined using REFMAC5 (ref. 32), and refined by cycles of manual rebuilding in COOT. Ligand coordinates and definitions were generated using the GlycoBioChem PRODRG2 server (<http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg>). The data and refinement statistics are available in Extended Data Table 1. **Native mass spectrometry.** Protein was buffer exchanged into 100 mM ammonium acetate using a 10-DG desalting column (BioRad). Native mass spectrometry experiments were carried out on a Synapt G2 instrument (Waters, Manchester, UK) with a nano-electrospray ionisation (nESI) source. Mass calibration was performed by a separate infusion of NaI cluster ions. Solutions were ionised through a positive potential applied to a platinum wire of thickness 0.125 mm (Goodfellow) inserted into a thin-walled glass capillary (inner diameter 0.9 mm, outer diameter 1.2 mm, World Precision Instruments, Stevenage, UK) that was pulled to a nESI tip in house with a Flaming/Brown micropipette puller (Sutter Instrument Co., Novato, CA). Fdc1 samples (5–10  $\mu\text{M}$ ) were sprayed from 100 mM ammonium acetate pH 6.8. Capillary voltage 1.6 kV, sample cone voltage 80–90 V, extractor cone voltage 3 V, backing pressure 5 mbar, source temperature 50 °C, trap gas flow 0.4–5  $\text{ml min}^{-1}$ . Data were processed using Masslynx V4.1 software and Originlab 9.0.

**Mass spectrometry analysis of the extracted Fdc1 cofactor.** Protein was buffer exchanged into deionised  $\text{H}_2\text{O}$  using a 10-DG column. 4  $\mu\text{l}$  1 M ammonium acetate was added to 200  $\mu\text{l}$  protein solution. The protein was precipitated and cofactor released with addition of 200  $\mu\text{l}$  acetonitrile and incubation at 70 °C for 5 min before centrifugation at 16,100g to remove the precipitate.

All solvents were of HPLC-MS Grade purity purchased from Sigma-Aldrich (Gillingham, UK). HPLC grade formic acid was purchased from Fisher Scientific (Loughborough, UK). Mass spectrometer calibration solution and chromatography columns were purchased from Thermo-Fisher Scientific (Hemel Hempstead, UK).

All UHPLC-MS work was carried out on a Thermo-Finnigan Orbitrap-LTQ XL hybrid mass spectrometer operated in negative ionization mode coupled to a Thermo Accela autosampler (Fisher Scientific, Bremen, Germany). Chromatographic separations were adapted from work carried out by Fu and co-workers<sup>34</sup> and performed on a Thermo Hypersil Gold 2.1  $\mu\text{m}$  C<sub>18</sub> column at a solvent flow-rate of 400  $\mu\text{l min}^{-1}$ . For initial profiling tests the column was eluted with 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The solvent composition during gradient elution was initiated with 5% (B) for 5 min and subsequently ramped to 95% (B) over 15 min, followed by a 5 min isocratic elution at 95% (B) before a return to 95% (A) held for further 5 min for column equilibration. All samples were maintained 4 °C within the autosampler refrigerator while the column was maintained at 50 °C within the autosampler oven. Mass calibration was carried out in accordance with the manufacturer's guidelines using

caffeine ( $20 \mu\text{g ml}^{-1}$ ), the tetrapeptide MRFA ( $1 \mu\text{g ml}^{-1}$ ) and Ultramark 1621 (0.001%) in an aqueous solution of acetonitrile (50%), methanol (25%) and acetic acid (1%). Acquisition settings for initial profiling were carried out at 60,000 resolution in centroid and ran at  $1 \mu\text{-scan per } 400 \text{ ms}$  in the  $100\text{--}1,000 \text{ m/z}$  range with source gasses set at sheath gas = 40 arbitrary units, aux gas = 5 arbitrary units, sweep gas = 5 arbitrary units. The ESI source voltage was set to 4.2 V, and capillary ion transfer tube temperature set at  $275^\circ\text{C}$ .

Mass fragmentation analysis was carried out with the same column chemistry, source settings and flow rate but with an isocratic solvent elution of 55% (A)/45% (B), an optimum composition determined for eluting the analyte of interest, based on the initial elution profile described above. Mass spectrometry source, samples storage/column was kept under identical conditions. Collision induced dissociation (CID) settings were set up to trap ions  $523 \text{ m/z}$  with an isolation width of  $1.0 \text{ m/z}$ , normalized collision energy of 35, activation Q of 0.250 and activation time of 30 ms.

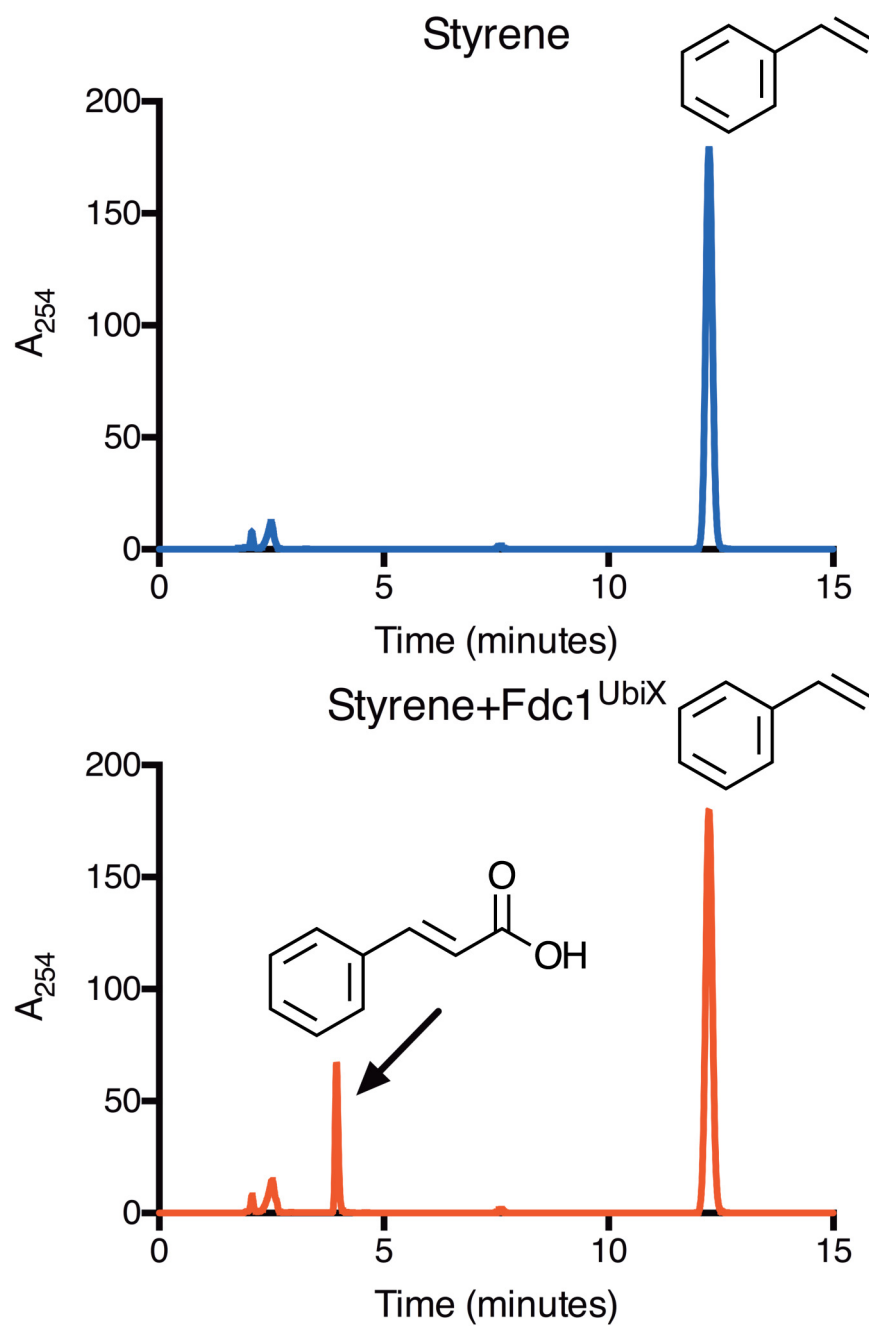
**DFT calculations.** Density function theory (DFT) models were optimized in the gas phase using the B3LYP/6-311++G(d,p) or BH&H/6-311++G(d,p) level of theory implemented in Gaussian 09 (ref. 35). Models are shown in Extended Data Fig. 7 and structural alignments to the crystal coordinates were performed using Swiss-PdbViewer version 4.1 (ref. 36). Harmonic vibrational frequencies calculated using normal mode analysis were used to confirm that optimized

geometries were always in local or global minima, and to calculate relative free energies.

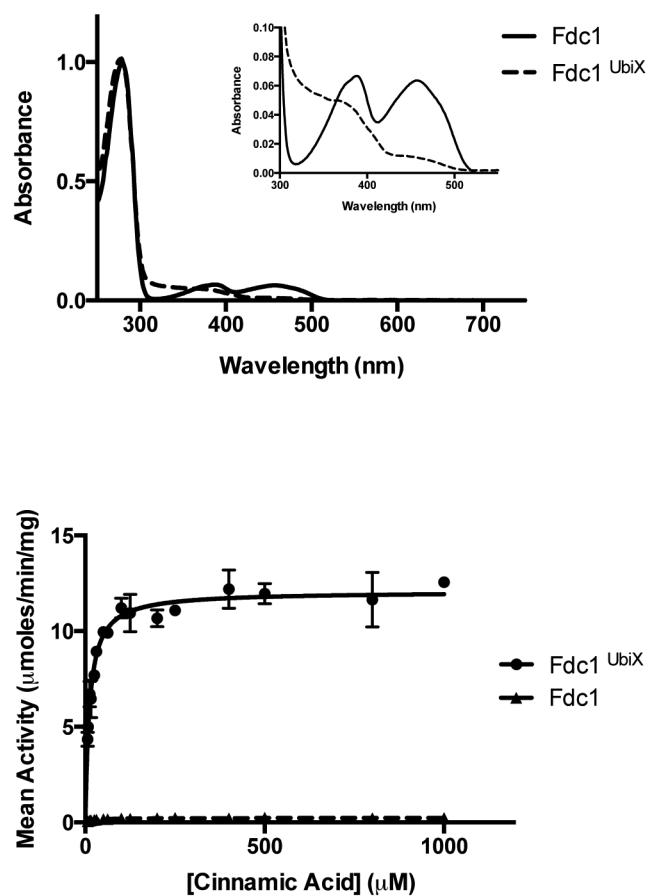
**Data reporting.** No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

30. Kalghatgi, K. K. & Subba Rao, P. V. Microbial L-phenylalanine ammonia-lyase. Purification, subunit structure and kinetic properties of the enzyme from *Rhizoctonia solani*. *Biochem. J.* **149**, 65–72 (1975).
31. de Villiers, M. M. & Bergh, J. J. Comparing HPLC and UV spectrophotometric analysis methods for determining the stability of sorbic acid in nonionic creams containing lactic acid. *Drug Dev. Ind. Pharm.* **26**, 539–547 (2000).
32. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
33. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
34. Fu, G. *et al.* Atomic-resolution structure of an N5 flavin adduct in D-arginine dehydrogenase. *Biochemistry* **50**, 6292–6294 (2011).
35. Frisch, X. M. J. *et al.* Gaussian 09 (Gaussian, Wallingford, CT, revision B.01, 2010).
36. Guex, Y. N. & Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
37. Waller, M. P., Robertazzi, A., Platts, J. A., Hibbs, D. E. & Williams, P. A. Hybrid density functional theory for pi-stacking interactions: application to benzenes, pyridines, and DNA bases. *J. Comput. Chem.* **27**, 491 (2006).

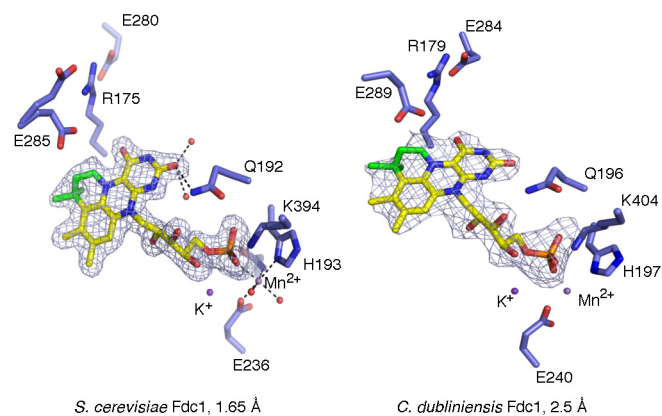




**Extended Data Figure 1** | HPLC chromatogram demonstrating enzymatic carboxylation of styrene by *A. niger* Fdc1<sup>UbiX</sup>. Chromatogram of a 10 mM styrene and saturating NaHCO<sub>3</sub> solution incubated in the absence (blue) and presence (red) of the *A. niger* Fdc1<sup>UbiX</sup> enzyme.

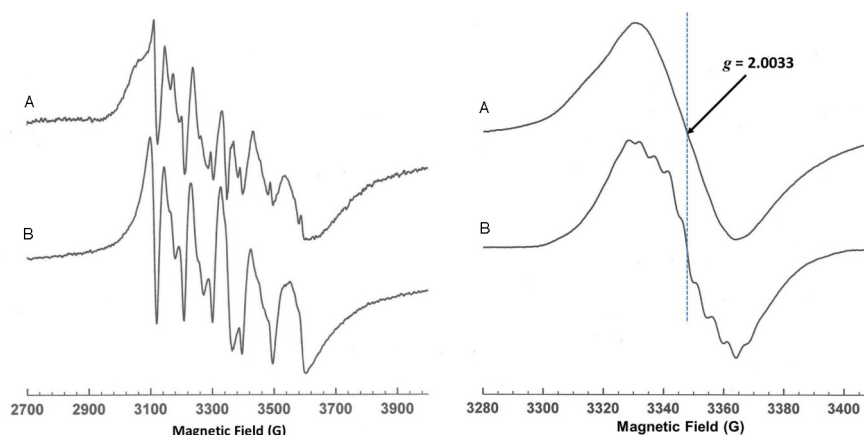


**Extended Data Figure 2** | *S. cerevisiae* Fdc1<sup>UbiX</sup> solution data. The top panel is a direct comparison to main Fig. 1b. In this case, the non-UbiX coexpressed Fdc1 binds FMN weakly. The bottom panel shows *S. cerevisiae* Fdc1<sup>UbiX</sup> steady-state kinetic data obtained with cinnamic acid. Error bars are s.e.m.,  $n = 3$



**Extended Data Figure 3 | A detailed view of the *S. cerevisiae* and *C. dubliniensis* Fdc1<sup>UbiX</sup> active site.** The omit map corresponding to the prFMN cofactor is shown in a blue mesh contoured at 3 $\sigma$ . Atoms derived from

dimethylallyl-monophosphate are shown in green. Both Fdc1<sup>UbiX</sup> crystals appeared colourless as observed for *A. niger* Fdc1<sup>UbiX</sup>, indicating negligible FMN binding.

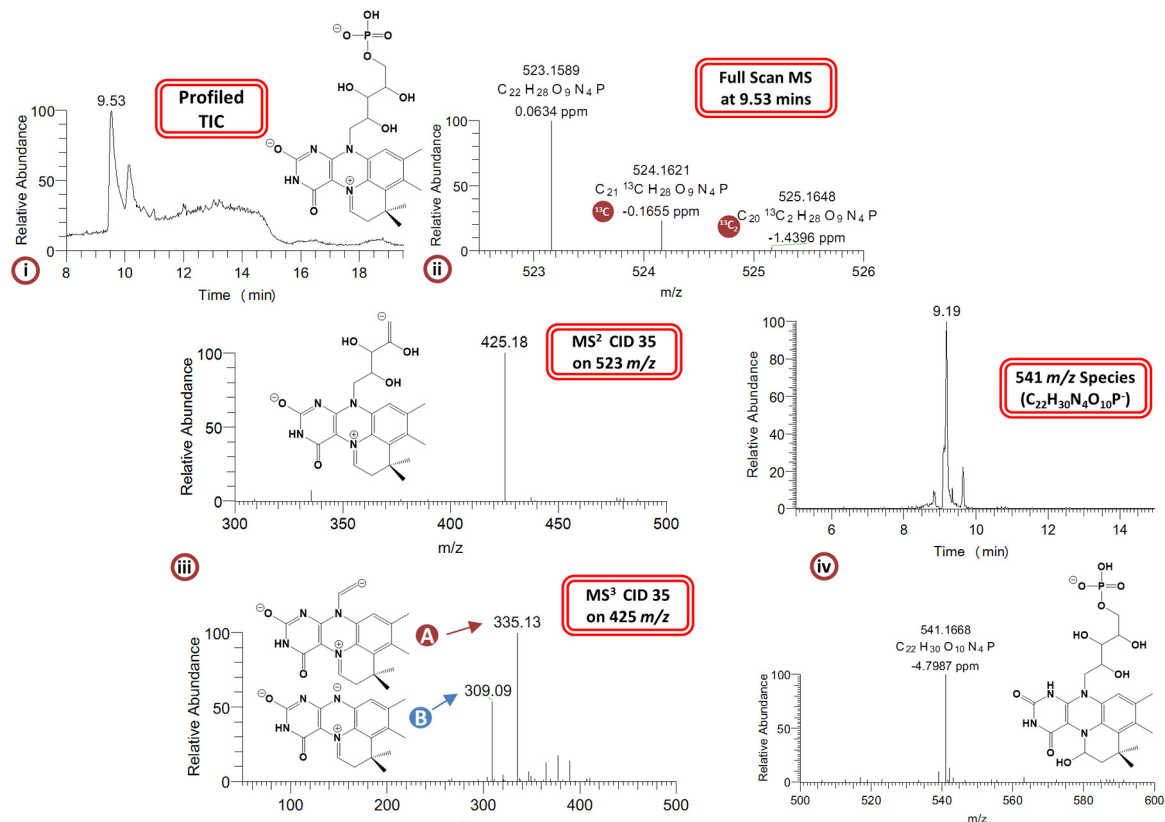


**Extended Data Figure 4 | EPR spectroscopy confirms  $\text{Mn}^{2+}$  binding.** Left panel shows X-band continuous wave frozen solution EPR spectra of (A)  $\text{Fdc1}^{\text{UbiX}}$  and (B)  $\text{Mn}(\text{H}_2\text{O})_6^{2+}$  showing the characteristic six line pattern arising from the  $m_s = \pm 1/2$  spin manifold of the  $S = 5/2$   $\text{Mn}^{2+}$  ion. The six line pattern reflects hyperfine coupling to the  $I = 5/2$   $^{55}\text{Mn}$  nucleus and is sensitive to the environment of the ion, thus the differences between A and B indicate binding of  $\text{Mn}^{2+}$  to the enzyme. Experimental conditions: microwave power 0.5 mW, modulation amplitude 7 G, temperature 20 K. The right panel shows X-band continuous wave frozen solution EPR spectra of (A)  $\text{Fdc1}^{\text{UbiX}}$  reduced using sodium cyanoborohydride ( $\text{NaBH}_3\text{CN}$ ) and subsequently exposed to air

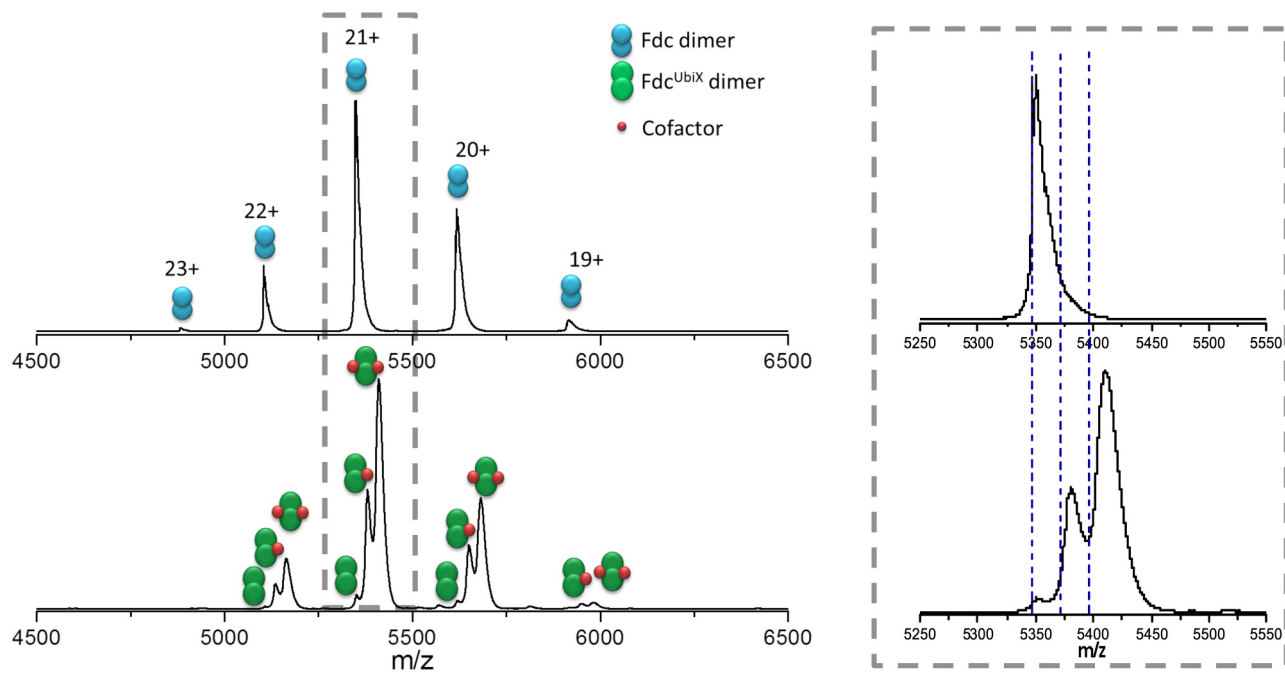
and (B) wild-type UbiX + DMAP reduced with dithionite and reoxidised with oxygen. Experimental conditions: microwave power 10  $\mu\text{W}$ , modulation amplitude 1.5 G, temperature 20 K. Reduction using  $\text{NaBH}_3\text{CN}$  of the modified cofactor formed in  $\text{Fdc1}^{\text{UbiX}}$  ( $\text{prFMN}^{\text{riminium}}$ ) gives rise to a radical on air oxidation with the same  $g$  value ( $g_{\text{av}}$ ) and line width as that formed on the modified flavin ( $\text{prFMN}$ ) in UbiX. However, the lack of the distinctive fine structure in A that is normally observed for the UbiX radical B, suggests heterogeneity in the  $\text{Fdc1}^{\text{UbiX}}$  radical or possibly a magnetic interaction with the nearby  $\text{Mn}^{2+}$  ion in  $\text{Fdc1}$ .



a



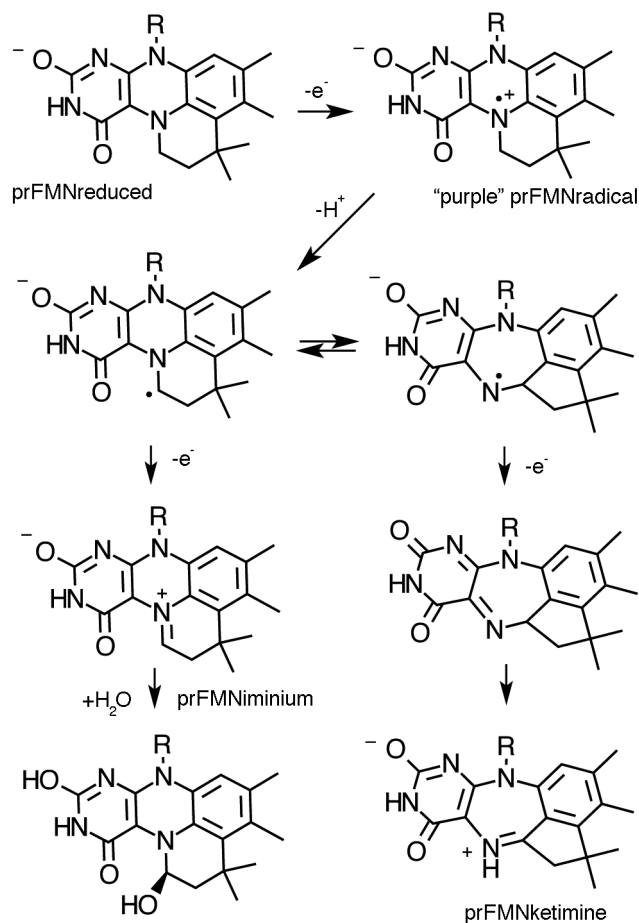
b



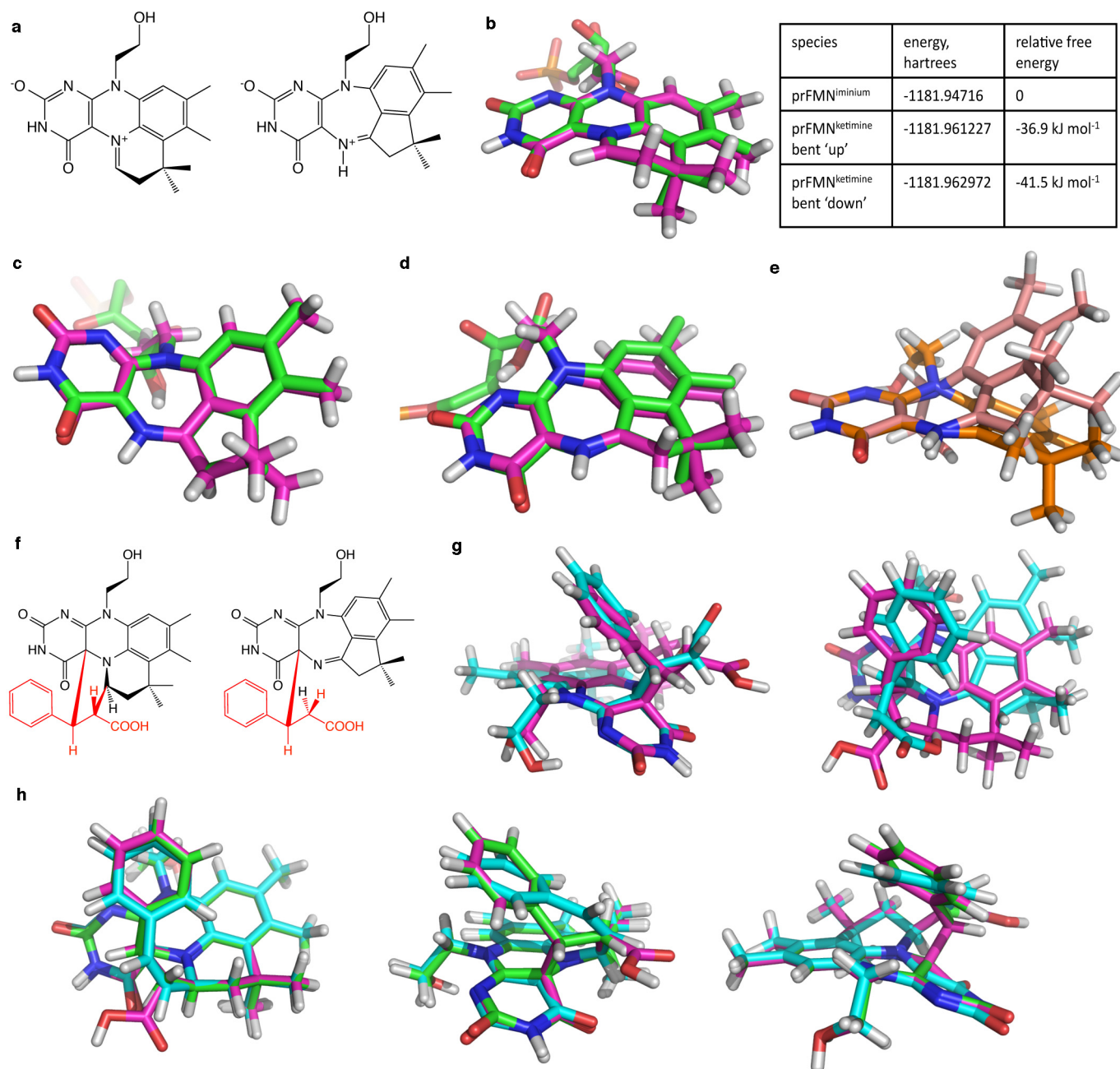
**Extended Data Figure 5 | Mass spectrometry of the Fdc1<sup>UbiX</sup> cofactor.**

**a,** (i) Structural elucidation of the Fdc1<sup>UbiX</sup> cofactor. Full scan total ion current (TIC) created under a gradient elution using H<sub>2</sub>O/acetonitrile both containing 0.1% formic acid indicating a major peak apex at 9.53 min with a 52/48 solvent composition. Also shown is the proposed structure of 523 *m/z*. Mass spectrum taken at major peak apex in (ii) (9.53 min) indicating an associated full-scan molecular ion peak with *m/z* = 523.1589 ( $M^+ = C_{22}H_{28}N_4O_9P$ ) at a resolution of 58,501 with a mass accuracy of 0.06 p.p.m. Also eluting alongside the target molecule are two isotopic variants containing <sup>13</sup>C and <sup>13</sup>C<sub>2</sub>. The <sup>13</sup>C peak is displaying a relative abundance of 22 to the 523 *m/z* peak that is in line with the number of carbon atoms contained within the structure. Fragmentation of the 523.16 *m/z* molecular ion peak (iii) in an automated data dependent manner using helium based-chemical induced dissociation (CID level 35) generated a spectral tree that indicates the removal of the phosphate head group at the MS<sup>2</sup> level generating 425.18 *m/z*. Subsequent MS<sup>3</sup> level activation (CID 35) on 425.18 *m/z* partially (A) or completely (B) removing the tail group from the newly formed four-ring system generating 335.13 *m/z* or 309.09 *m/z*, respectively. The presence of a hydrolysed species with 541 *m/z* is also reported (iv). **b,** Native mass spectra of Fdc1 (top) and Fdc1<sup>UbiX</sup> (bottom). Fdc1 presents in charge states 19+ to 23+, while Fdc1<sup>UbiX</sup> presents in charge states 19+ to

22+. Right-hand spectrum; an enlarged view of the 21+ charge state. The predicted masses are shown by blue dashed lines. The spectrum of Fdc1<sup>UbiX</sup> shows that approximately two-thirds of the ions have two non-covalently bound cofactors, approximately one-third have one non-covalently bound cofactor and there is a small amount with no cofactor bound. Fdc1 contains no cofactor. The measured mass of the Fdc1 dimer is 112,265 Da (predicted mass from sequence is 112,270 Da). The measured mass for the *apo* form of Fdc1<sup>UbiX</sup> is 112 345 Da, slightly higher than for Fdc1 which is attributed to an increased retention of salt. The mass difference of +80 Da corresponds to the mass of two potassium adducts. For the Fdc1<sup>UbiX</sup> species with one bound cofactor, the measured mass (112,968 Da) is 178 Da higher than predicted. The predicted mass corresponds to the left hand side of the peak, which is the protein plus cofactor with no extra salt retained. The extra mass could be attributed to 2 Mn<sup>2+</sup> ions and 2 K<sup>+</sup> ions. The Fdc1<sup>UbiX</sup> bound to two cofactors has a measured mass of 113,583 Da, which is 268 Da larger than expected. Again, however, the predicted mass corresponds to the left-hand side of the peak. These spectra indicate that the protein dimer carries either 1 or 2 cofactors of 525 Da, along with a variety of other salt adducts. The extent of adduction is higher for Fdc1<sup>UbiX</sup> and increases with bound cofactors, indicating that the addition of the cofactor recruits counter ions.



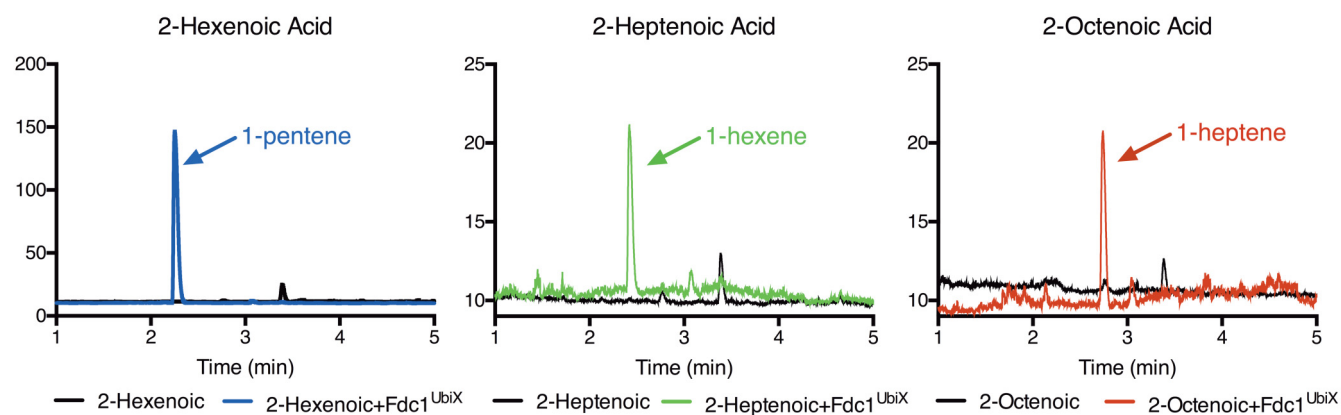
**Extended Data Figure 6 | A proposed mechanism for cofactor maturation in Fdc1<sup>UbiX</sup> through oxidation.** We propose the prFMN<sup>reduced</sup> cofactor produced by UbiX is bound by *apo*-Fdc1 and oxidized in a stepwise manner. Although the initial radical species resembles that observed during non-physiological oxidation of the prFMN<sup>reduced</sup>-UbiX product complex<sup>10</sup>, we propose that in the Fdc1 enzyme proton abstraction from the prFMN<sup>radical</sup> C1' leads to a distinct radical species. The latter can either be oxidized further to the corresponding prFMN<sup>iminium</sup> or, via a radical based isomerization process, to form the central seven-membered ring ultimately leading to prFMN<sup>ketimine</sup>.



**Extended Data Figure 7 | DFT models of the proposed prFMN<sup>iminium</sup> and prFMN<sup>ketimine</sup> isomers and their substrate adducts.** **a**, Chemical structures of the prFMN<sup>iminium</sup> and prFMN<sup>ketimine</sup> models. DFT models (pink carbons) are overlaid with crystal coordinates (green carbons) of prFMN<sup>ketimine</sup> (**b**) and the two butterfly bent conformations of prFMN<sup>ketimine</sup> (**c**, **d**). **c**, **d**, The substrate-free (bent 'up') prFMN<sup>ketimine</sup> is shown in **c** and the more planar substrate-bound (bent 'down') in **d**. For comparison of the extent of the butterfly bending of prFMN<sup>ketimine</sup>, the two DFT models in **c** and **d** are aligned over the four-ring nitrogen atoms and overlaid in **e**. **f**, Chemical structures of the proposed initial prFMN<sup>iminium</sup> and prFMN<sup>ketimine</sup> substrate adducts with the cinnamic acid substrate highlighted in red. The DFT-optimized structures of these species are overlaid in **g** with the prFMN<sup>iminium</sup> species shown with pink carbons and the prFMN<sup>ketimine</sup> species with teal carbons. Three projections of overlaid DFT models of the prFMN<sup>iminium</sup> species with substrate bound (pink carbons; reproduced from **g**), after substrate decarboxylation (teal carbons; substrate

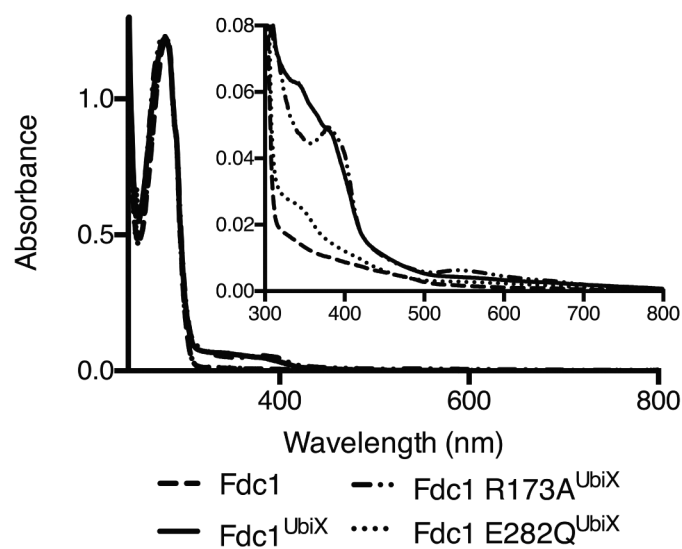
double bond is *cis*) and upon protonation of the substrate  $\beta$ -carbon (green carbons) are shown in **h**). Note that the substrate carboxylate was artificially protonated in these models to maintain charge neutrality. Models were geometry optimized in the gas phase using either the B3LYP/6-311++G(d,p) (panels **a–e**) or BH&H/6-311++G(d,p) (panels **f–h**) level of theory. BH&H was chosen over B3LYP for the substrate adducts as BH&H has been shown to better describe  $\pi$ -stacking interactions, which are likely to occur between the modified isoalloxazine and substrate phenyl moieties<sup>37</sup>. Harmonic vibrational frequencies calculated using normal mode analysis were used to confirm that optimized geometries of all species were in local or global minima. Absolute energies and relative free energies of the substrate-free species, determined from the normal mode calculations, are given in the table, top right. Cartesian coordinates of the optimized structures are given in the Supplementary Information.





**Extended Data Figure 8 | Fdc1 decarboxylase activity with simple acrylic acids.** Gas chromatogram showing products formed from a solution of respectively 10 mM 2-hexenoic (1-pentene, blue; enzyme free control in black),

10 mM 2-heptenoic (1-hexene, green; enzyme free control in black) or 10 mM 2-octenoic acid (1-heptene, red; enzyme free control in black). Identification of 1-alkenes was based on known standards.



**Extended Data Figure 9** | **UV-visible spectral properties of Fdc1 variants.** UV-visible spectra of Fdc1 (614  $\mu$ M), Fdc1<sup>UbiX</sup> (492  $\mu$ M), Fdc1 R173A<sup>UbiX</sup> (749  $\mu$ M) and Fdc1 E282Q<sup>UbiX</sup> (171  $\mu$ M) normalised on the  $A_{280}$  peak. Inset, close up of the additional spectral features present in the 300–500 nm region.

Extended Data Table 1 | Data collection and refinement statistics

|   | <i>A. niger</i> Fdc1<br>(only<br>prFMN <sup>iminium</sup><br>present) | <i>A. niger</i> Fdc1<br>(prFMN <sup>iminium</sup><br>+<br>prFMN <sup>ketimine</sup> ) | <i>A. niger</i> Fdc1<br>+ alpha-<br>methyl<br>cinnamic acid | <i>A. niger</i> Fdc1<br>+ penta-fluoro<br>cinnamic acid | <i>A. niger</i> Fdc1<br>+ alpha-fluoro<br>cinnamic acid | <i>A. niger</i> Fdc1<br>+ phenyl<br>pyruvate | <i>A. niger</i> Fdc1<br>+ 4-vinyl<br>guaiacol | <i>S. cerevisiae</i><br>Fdc1 | <i>C. dubliniensis</i><br>Fdc1 |
|---|---|---|---|---|---|--|---|------------------------------|--------------------------------|
| PDB code  | 4ZA4  | 4ZA5  | 4ZA7  | 4ZA8  | 4ZAB  | 4ZA9   | 4ZAA  | 4ZAC                         | 4ZAD                           |
| <b>Data collection</b>                                |   |   |   |   |   |  |   |                              |                                |
| Space group   | P 2 <sub>1</sub> 2 <sub>1</sub> 2                                     | P 2 <sub>1</sub> 2 <sub>1</sub> 2   | P 2 <sub>1</sub> 2 <sub>1</sub> 2                           | P 2 <sub>1</sub> 2 <sub>1</sub> 2                       | P 2 <sub>1</sub> 2 <sub>1</sub> 2                       | P 2 <sub>1</sub> 2 <sub>1</sub> 2            | P 2 <sub>1</sub> 2 <sub>1</sub> 2             | P 2 <sub>1</sub>             | P 2 <sub>1</sub>               |
| Cell dimensions<br><i>a</i> , <i>b</i> , <i>c</i> (Å) | 95.99<br>63.93<br>87.63   | 96.02<br>63.79<br>87.72   | 95.93<br>64.16<br>87.73                                     | 95.91<br>64.20<br>87.79                                 | 95.91<br>64.20<br>87.79                                 | 96.05<br>63.85<br>87.68                      | 96.59<br>64.47<br>87.89                       | 114.41<br>96.18<br>116.64    | 91.97<br>64.55<br>96.05        |
| $\beta$ (°)   |   |   |   |   |   |  |   | 96.58                        | 91.07                          |
| Resolution (Å)  | 33.82-1.22<br>(1.25-1.22)   | 48.01-1.10<br>(1.13-1.10)   | 87.73-1.1<br>(1.13-1.10)                                    | 64.81-1.06<br>(1.09-1.06)                               | 32.38-1.16<br>(1.19-1.16)                               | 63.85-1.01<br>(1.04-1.01)                    | 94.19-1.242<br>(1.274-<br>1.242)              | 54.19-1.65<br>(1.69-1.65)    | 67.04-2.46<br>(2.52-2.46)      |
| $R_{\text{pim}}$                                      | 3.7 (35.4)  | 3.2 (37.7)  | 3.5 (30.9)  | 3.9 (33.9)  | 3.8 (36.1)  | 2.7 (31.4)                                   | 3.1 (34.4)                                    | 3.9 (39.2)                   | 9.1 (48.0)                     |
| $I/\sigma I$  | 11.2 (2.0)  | 12.6 (2.1)  | 11.20 (2.4)   | 12.8 (2.5)  | 12.0 (2.3)  | 14.3 (2.2)                                   | 11.5 (2.0)                                    | 11.0 (2.0)                   | 6.7 (1.7)                      |
| Completeness (%)                                      | 99.96<br>(99.99)  | 99.97<br>(99.99)  | 98.0(97.8)  | 99.96<br>(99.98)  | 99.9 (99.9)   | 99.97<br>(99.99)                             | 98.9 (94.4)                                   | 98.8 (99.7)                  | 99.8 (99.8)                    |
| Redundancy  | 6.1 (5.8)   | 5.5 (3.6)   | 4.2 (2.7)   | 5.2 (3.6)   | 4.2 (2.8)   | 4.0 (2.2)                                    | 6.5 (6.0)                                     | 3.7 (3.8)                    | 3.4 (3.4)                      |
| <b>Refinement</b>                                     |   |   |   |   |   |  |   |                              |                                |
| Resolution (Å)  | 33.82-1.22<br>(1.25-1.22)   | 48.01-1.10<br>(1.13-1.10)   | 30.05-1.10<br>(1.13-1.10)                                   | 64.81-1.06<br>(1.09-1.06)                               | 32.38-1.16<br>(1.19-1.16)                               | 63.85-1.01<br>(1.04-1.01)                    | 94.19-1.25<br>(1.274-<br>1.242)               | 54.19-1.65<br>(1.69-1.65)    | 67.04-2.46<br>(2.52-2.46)      |
| No. reflections                                       | 152260<br>(8041)  | 206025<br>(10689)   | 203456<br>(10579)   | 231449<br>(12068)                                       | 175535<br>(9183)  | 264342<br>(13893)                            | 146799<br>(7752)                              | 282233<br>(15010)            | 39186<br>(1956)                |
| $R_{\text{work}}/R_{\text{free}}$                     | 14.5/16.1<br>(23.6/26.4)  | 14.5/15.6<br>(25.9/26.9)  | 13.2/14.1<br>(20.7/21.2)                                    | 14.79/16.61<br>(23.1/26.9)                              | 14.6/15.4<br>(23.0/23.5)                                | 12.99/14.50<br>(24.6/26.0)                   | 14.34/16.32<br>(25.4/30.0)                    | 16.44/19.23<br>(31.6/34.8)   | 20.66/25.5<br>(35.2/44.2)      |
| No. atoms   | 4595  | 4692  | 4649  | 4665  | 4694  | 4554   | 4661  | 17693                        | 8213                           |
| B-factors   | 9.81  | 10.71   | 10.25   | 10.321  | 9.83  | 9.38   | 11.07   | 24.8                         | 26.1                           |
| R.m.s deviations                                      |   |   |   |   |   |  |   |                              |                                |
| Bond lengths (Å)                                      | 0.028   | 0.028   | 0.028   | 0.029   | 0.031   | 0.029  | 0.026   | 0.032                        | 0.017                          |
| Bond angles (°)                                       | 2.384   | 2.485   | 2.469   | 2.429   | 2.622   | 2.250  | 2.304   | 2.467                        | 1.726                          |

# UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis

Mark D. White<sup>1</sup>, Karl A. P. Payne<sup>1</sup>, Karl Fisher<sup>1</sup>, Stephen A. Marshall<sup>1</sup>, David Parker<sup>2</sup>, Nicholas J. W. Rattray<sup>1</sup>, Drupad K. Trivedi<sup>1</sup>, Royston Goodacre<sup>3</sup>, Stephen E. J. Rigby<sup>1</sup>, Nigel S. Scrutton<sup>1</sup>, Sam Hay<sup>1</sup> & David Leys<sup>1</sup>

Ubiquinone (also known as coenzyme Q) is a ubiquitous lipid-soluble redox cofactor that is an essential component of electron transfer chains<sup>1</sup>. Eleven genes have been implicated in bacterial ubiquinone biosynthesis, including *ubiX* and *ubiD*, which are responsible for decarboxylation of the 3-octaprenyl-4-hydroxybenzoate precursor<sup>2</sup>. Despite structural and biochemical characterization of UbiX as a flavin mononucleotide (FMN)-binding protein, no decarboxylase activity has been detected<sup>3,4</sup>. Here we report that UbiX produces a novel flavin-derived cofactor required for the decarboxylase activity of UbiD<sup>5</sup>. UbiX acts as a flavin prenyltransferase, linking a dimethylallyl moiety to the flavin N5 and C6 atoms. This adds a fourth non-aromatic ring to the flavin isoalloxazine group. In contrast to other prenyltransferases<sup>6,7</sup>, UbiX is metal-independent and requires dimethylallyl-monophosphate as substrate. Kinetic crystallography reveals that the prenyltransferase mechanism of UbiX resembles that of the terpene synthases<sup>8</sup>. The active site environment is dominated by  $\pi$  systems, which assist phosphate-C1' bond breakage following FMN reduction, leading to formation of the N5-C1' bond. UbiX then acts as a chaperone for adduct reorientation, via transient carbocation species, leading ultimately to formation of the dimethylallyl C3'-C6 bond. Our findings establish the mechanism for formation of a new flavin-derived cofactor, extending both flavin and terpenoid biochemical repertoires.

Flavin is a common cofactor responsible for highly versatile (redox) chemistry<sup>9</sup>, its properties modified or fine-tuned by the protein scaffold, occasionally through covalent attachment to the protein matrix via the C6 or C8 atoms of the isoalloxazine dimethylbenzene ring<sup>10</sup>. We have shown that the reversible decarboxylases belonging to the UbiD or the homologous Fdc enzyme family require a heavily modified FMN cofactor<sup>5</sup>. In these enzymes, the FMN is linked to a C<sub>5</sub>-alkyl group via N5-C1' and C6-C3' linkages, adding a fourth (non-aromatic) ring to the isoalloxazine ring system. Knockout of *ubiD/fdc1* or the associated *ubiX* (or the homologous *pad1*) genes leads to similar phenotypes, and the latter have been shown to encode for FMN-binding proteins for which no decarboxylase activity has been detected *in vitro*<sup>2-4,11</sup>. Recently it was found that UbiX/Pad1 is responsible for activation of *Saccharomyces cerevisiae* Fdc1<sup>12</sup>, suggesting that UbiX/Pad1 is responsible for the synthesis of the UbiD/Fdc1 cofactor. The observed FMN modification by a C<sub>5</sub>-unit suggests this occurs through prenyltransferase activity (Fig. 1a).

We tested this hypothesis by incubating *Pseudomonas aeruginosa* UbiX<sup>13</sup> with oxidized FMN and the universal isoprene precursors dimethylallyl-pyrophosphate (DMAPP) or isopentenyl-pyrophosphate (IPP). However, spectroscopic evidence for formation of a ternary isoprene precursor-FMN-UbiX complex or of covalent FMN modification could not be obtained. By contrast, perturbation in the FMN ultraviolet (UV)-visible spectrum was observed in the presence of dimethylallyl-monophosphate (DMP), leading to an apparent  $K_d$  of  $12.0 \pm 0.4 \mu\text{M}$  (Fig. 1b). Although the minor spectral perturbation

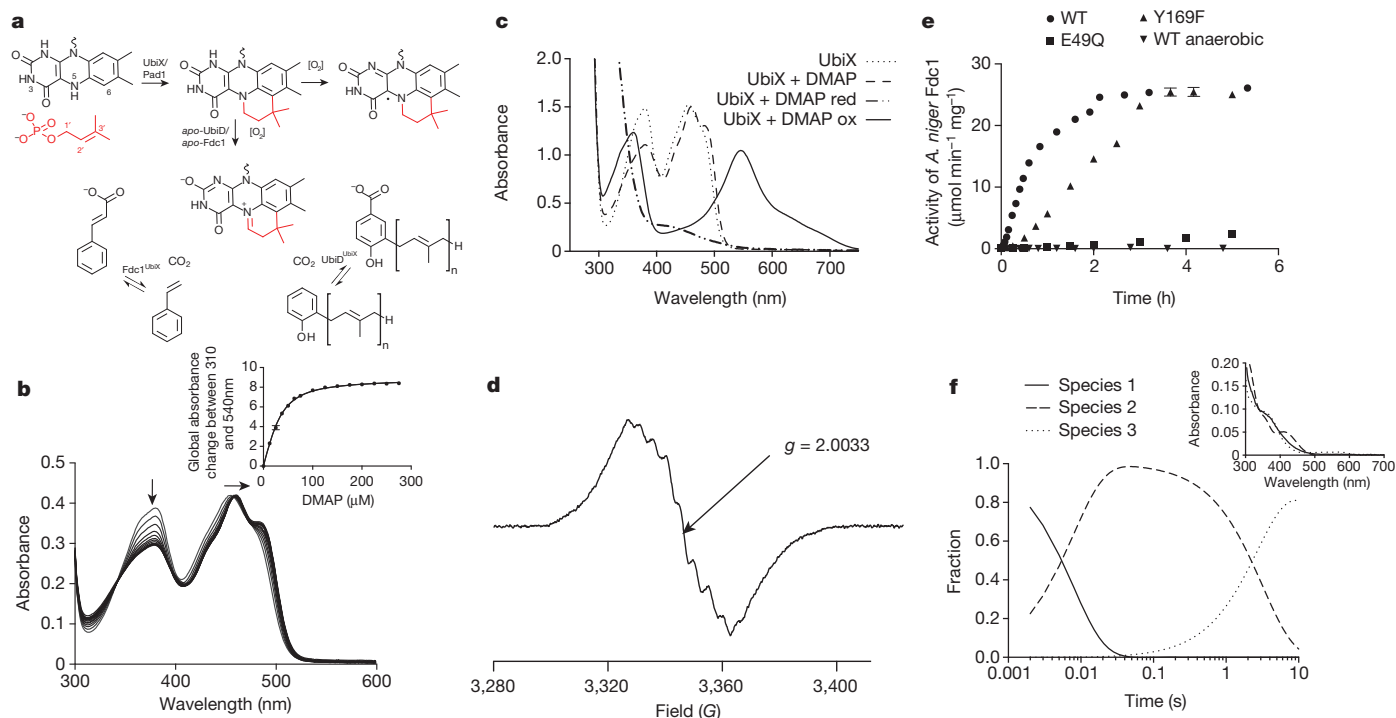
is indicative of binding close to the FMN isoalloxazine, it does not reflect FMN modification. However, reduction of the ternary FMN-DMP-UbiX complex using sodium dithionite followed by re-oxidation under aerobic conditions leads to formation of a stable purple-coloured intermediate (Fig. 1c). Ultra-high-performance liquid chromatography coupled with high-resolution mass spectrometry of extracts from both the reduced and re-oxidized ternary FMN-DMP-UbiX complex reveal a molecular species with mass corresponding to the reduced (prenylated-FMN<sup>reduced</sup> or prFMN<sup>reduced</sup>) or radical (prFMN<sup>radical</sup>) form of the UbiD/Fdc1 cofactor, respectively (Extended Data Fig. 1) and electron paramagnetic resonance (EPR) spectroscopy confirms the formation of a radical in ~95% yield (relative to FMN) in the reoxidized ternary complex (Fig. 1d). Furthermore, the radical species has EPR properties consistent with that of a N5, C6 alkylated flavin semiquinone (Extended Data Figs 2-4). Activity of *Aspergillus niger* apo-Fdc1 could be reconstituted *in vitro* following anaerobic incubation with prFMN<sup>reduced</sup>-UbiX. Decarboxylase activity could only be observed after exposure of the reconstituted Fdc1 to oxygen (Fig. 1e), but appeared independent of the presence of UbiX (Extended Data Fig. 5a). No activity could be observed when incubating apo-Fdc1 with the prFMN<sup>radical</sup>-UbiX. This suggests only the prFMN<sup>reduced</sup> form can be correctly oxidized by UbiD/Fdc1 to the corresponding reactive N5-iminium species (prFMN<sup>iminium</sup>).

Stopped-flow experiments mixing reduced FMNH<sub>2</sub>-UbiX with DMAP under anaerobic conditions reveal transient formation of a distinct spectral species (species 2 in Fig. 1f) before formation of the prFMN<sup>reduced</sup> product. The rate of species 2 formation showed a hyperbolic dependence on DMAP concentration with apparent  $K_d = 730 \pm 70 \mu\text{M}$  and limiting rate of  $k_f = 177 \pm 7 \text{ s}^{-1}$  (Extended Data Fig. 5b, c). The decay was independent of DMAP concentration and occurred with  $k_d = 0.316 \pm 0.002 \text{ s}^{-1}$ . Stopped-flow experiments mixing reduced prFMN<sup>reduced</sup>-UbiX with oxygen reveal that the rate of formation of the purple-coloured end product is linearly dependent on oxygen concentration (Extended Data Fig. 5d, e).

We obtained the 1.8 Å crystal structure of *P. aeruginosa* UbiX in complex with oxidized FMN and DMAP (Fig. 2a, b). The DMAP substrate is located directly above the FMN isoalloxazine *re* face, with the dimethylallyl moiety sandwiched between the FMN and A89-S90. The dimethylallyl group is furthermore surrounded by aromatic ring systems of W84, Y169 and W200 that, together with the FMN dimethylbenzene moiety, resemble the  $\pi$ -cage found in other prenyltransferases or terpene synthases<sup>6-8</sup>. The phosphate moiety is bound by several cationic residues, including conserved residues R122, K129, R139 and R185 in addition to polar interactions with S90, the backbone amide of G91, E140 and Y169 (Extended Data Fig. 6). The short N5 to substrate C1' distance (3.3 Å) and relative small N10-N5-C1' angle of 98° are compatible with those observed for other flavin-substrate complexes<sup>9,14</sup>. No evidence of direct nucleophilic/oxidative attack of the oxidized FMN N5 was observed. Substrate complex

<sup>1</sup>Centre for Synthetic Biology of Fine and Speciality Chemicals, Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, UK. <sup>2</sup>Innovation/Biobdomain, Shell International Exploration and Production, Westhollow Technology Center, 3333 Highway 6 South, Houston, Texas 77082-3101, USA.





**Figure 1** | *P. aeruginosa* UbiX solutions studies. **a**, Schematic overview of the proposed UbiX reaction. The N5–C6 prenylated FMNH<sub>2</sub> product (prFMN<sup>reduced</sup>) undergoes (probably non-physiological) oxidation to a radical species (prFMN<sup>radical</sup>) in the presence of oxygen (see **d**). In the presence of apo-UbiD or apo-Fdc1, we propose that the UbiX product is oxidized to the UbiD/Fdc1 prFMN<sup>iminium</sup> cofactor (see **e**). **b**, Titration of oxidized FMN–UbiX with DMAP. Grey lines represent individual scans at increasing DMAP concentrations with the black line representing the final spectrum obtained at saturation. The binding curve is derived from the global absorbance change in the 310–450 nm range. **c**, UV-visible spectra of UbiX–FMN obtained during

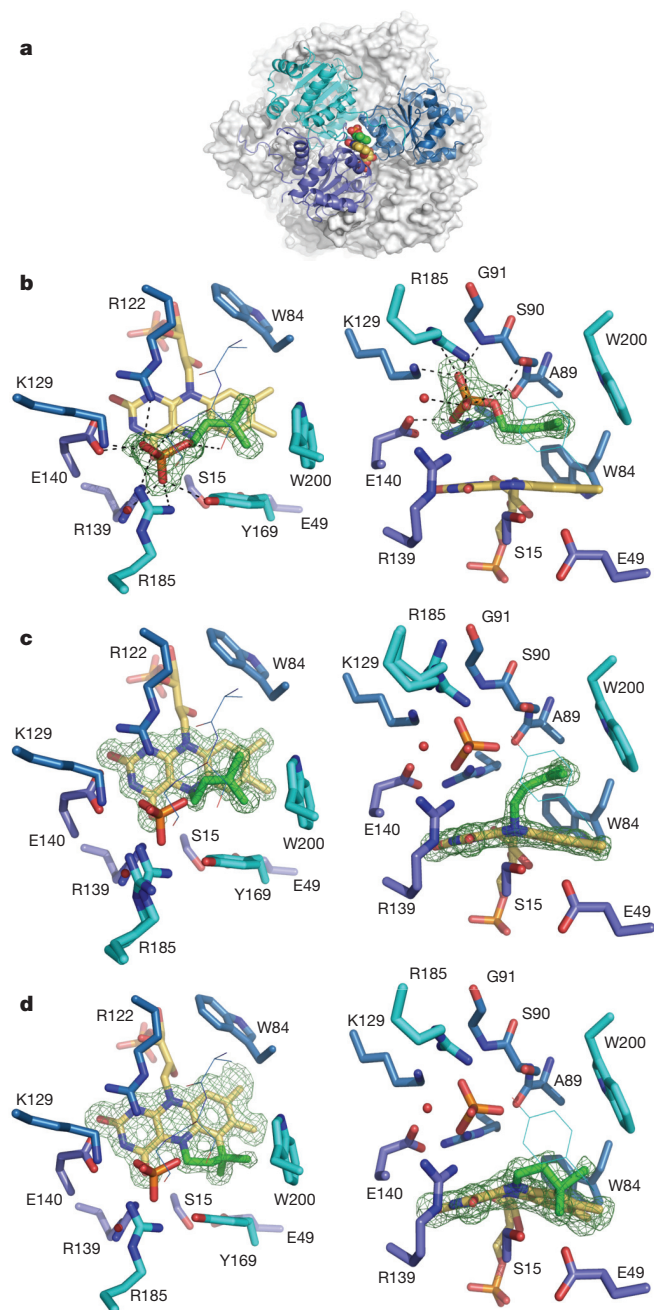
redox cycling in presence of DMAP. **d**, EPR spectrum of the UbiX–prFMN<sup>radical</sup> complex. **e**, Reconstitution of *A. niger* apo-Fdc1 decarboxylase activity by incubation with wild-type and variant UbiX enzymes in presence of DMAP and FMNH<sub>2</sub> followed by oxygen exposure. No activity can be observed under anaerobic conditions. **f**, Singular value decomposition of rapid-scan stopped-flow spectrophotometric data following mixing of wild-type UbiX–FMNH<sub>2</sub> with DMAP. The spectral species identified can tentatively be assigned to (1) ternary UbiX–FMNH<sub>2</sub>–DMAP complex; (2) an intermediate covalent adduct formed between FMNH<sub>2</sub> and DMAP and (3) the UbiX–prFMN<sup>reduced</sup> product.

crystals that are transiently reduced with sodium dithionite under aerobic conditions turned purple upon oxidation, closely resembling solution properties. The 1.6 Å crystal structure of a purple-coloured crystal reveals formation of the four-ring prFMN<sup>radical</sup> (Fig. 2d). Although product formation has little effect on the active site structure, the presence of the fourth ring is accompanied by a slight rotation of the FMN at the N3 atom. This accommodates the extended isalloxazine ring system while keeping the hydrophilic network involved in binding the isalloxazine ring intact. These observations establish UbiX crystals are catalytically competent, but do not reveal the order in which both N5–C1' and C6–C3' linkages are made.

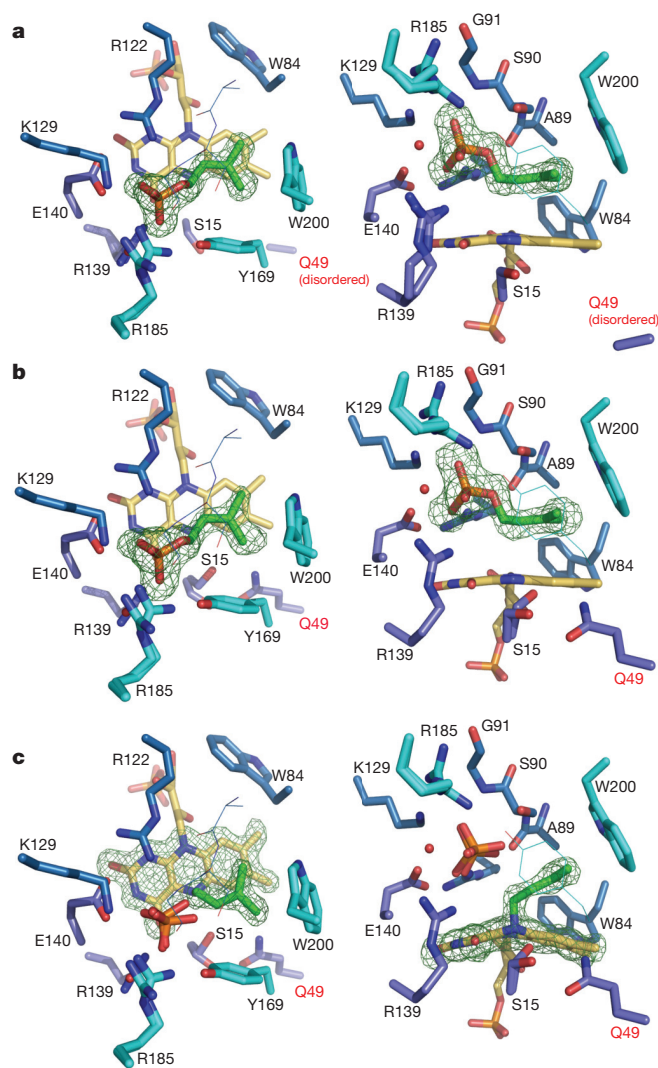
We used kinetic crystallography to provide detailed insights into the UbiX mechanism, taking advantage of the fact that a relatively long-lived intermediate can be observed in solution studies (Fig. 1f). Reduction of the enzyme–substrate complex by sodium dithionite followed by rapid-freezing and structure determination to 1.4 Å reveals that formation of an unusual N5–C1' dimethylallyl adduct occurs within 1–5 s, establishing that N5–C1' bond formation occurs first (Fig. 2c). The N5–C1' adduct adopts a significant butterfly conformation, and the N5 is clearly *sp*<sup>3</sup> hybridized. Although S15 and E49 are located in proximity of the N5, a (transient) reorientation of both side chains is required to establish a hydrogen-bonding network with the N5. The 1.9 Å structure of substrate complex crystals frozen 20–30 s following reduction reveals formation of the prFMN<sup>reduced</sup> product (Extended Data Fig. 7). Longer incubation times lead to formation of the prFMN radical species as indicated by a gradual purple colouration of crystals. No other intermediates were seen to accumulate within the wild-type crystals, in agreement with our stopped-flow solution data (Fig. 1f).

In order to determine whether S15 and/or E49 are involved in N5 deprotonation and/or formation of the C6–C3' linkage, we created an E49Q variant. The mutation severely affects but does not abolish *in vitro* Fdc1 activation (Fig. 1e). Although the oxidized substrate E49Q complex is similar in structure to the corresponding wild-type structure (Fig. 3a), a distinct intermediate species was seen following reduction and rapid freezing within 1–5 s (Fig. 3b). Although the latter structure clearly contains a reduced FMNH<sub>2</sub>, no N5–C1' bond has formed, in contrast to the corresponding wild-type intermediate structure obtained within 1–5 s following reduction (Fig. 2c). Furthermore, both S15 and Q49 establish a hydrogen-bonding network with the N5. This suggests N5 deprotonation through S15 and E49 is linked to N5–C1' bond formation, a process largely rendered ineffective through the E49Q mutation (Fig. 4, species I<sup>reduced</sup>). Structures of the E49Q variant derived for crystals frozen 0.5–10 min after reduction reveal that accumulation of an N5–C1' dimethylallyl adduct does occur (Fig. 3c), albeit distinct in conformation from that observed in the 1–5 s wild-type structure (Fig. 2c).

In the E49Q N5–C1' dimethylallyl adduct, the dimethylallyl substrate-derived moiety has undergone a 180° rotation when compared to the wild-type N5-adduct species, along with a small rotation of the isalloxazine ring (Figs 3c and 4, species IVa). Furthermore, both S15 and Q49 form a hydrogen-bonding network with N5, with the S15–N5 hydrogen bond maintaining the N5 in an *sp*<sup>3</sup> state. This series of conformational changes achieves two objectives: a proton relay network is established between the N5 and solvent via S15 and E49 and additional space is created between the isalloxazine dimethylbenzene moiety and Y169 to allow for formation of the fourth ring.

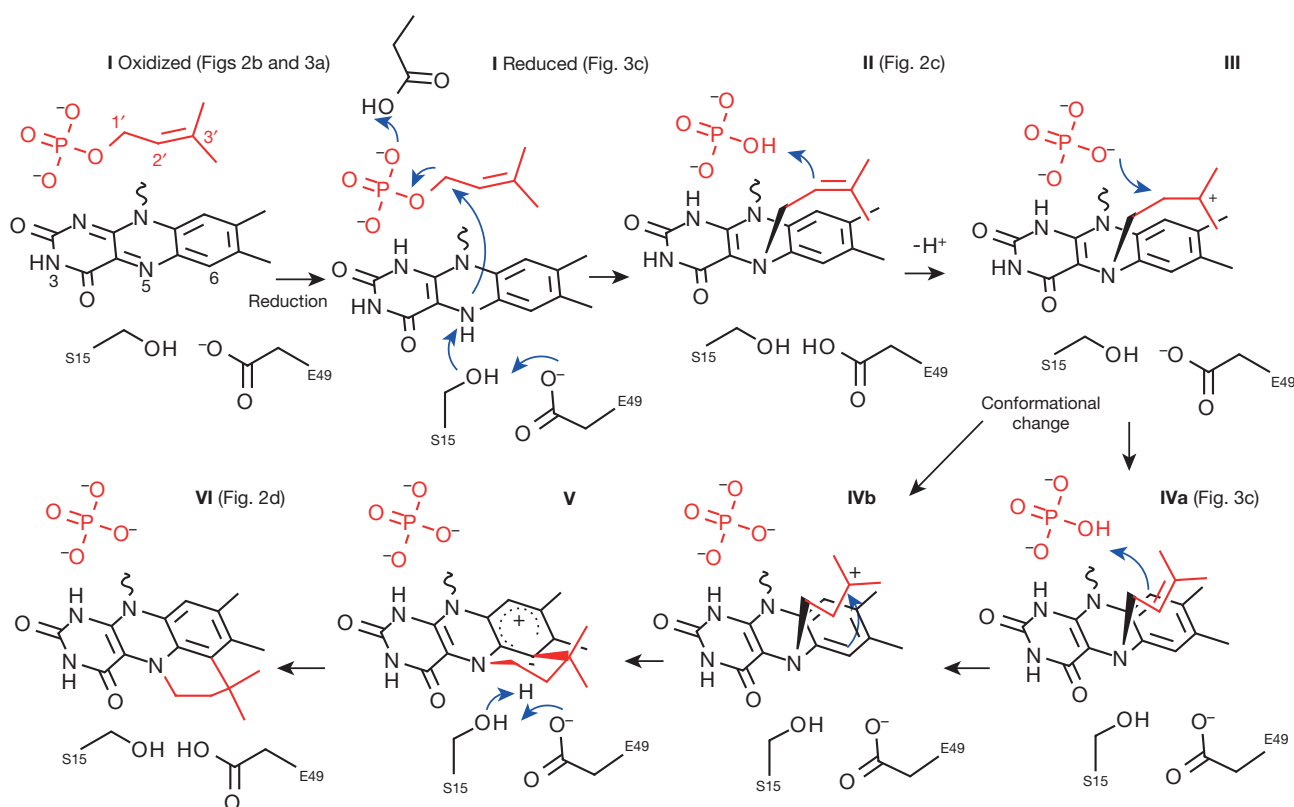


**Figure 2 | Crystal structures of *P. aeruginosa* UbiX.** **a**, Overview of the dodecameric UbiX structure, with 3 individual monomers (coloured in cyan, blue and purple, respectively) contributing to a single active site shown in cartoon representation while the remainder (in grey) is shown in surface representation. The bound FMN and DMAP substrates are shown in atom-coloured spheres. **b**, Detailed view of the oxidized UbiX-FMN-DMAP complex in two orientations related by a 90 degree rotation along the horizontal axis. Individual amino acids contributing to the active site structure are shown in atom coloured sticks (carbons colour coded as in Fig. 2a). Residues positioned in front of the FMN-DMAP substrates are shown as thin lines for clarity. The omit map for the DMAP substrate is shown as a green mesh, contoured at  $4\sigma$ . Hydrogen-bonding networks established with the phosphate moiety are shown by dotted lines. **c**, Detailed view of the N5-C1' alkyl adduct species (in two orientations related by a 90 degree rotation along the horizontal axis) obtained through rapid flash-cooling following reduction. The omit map for the N5-C1' adduct is shown as a green mesh, contoured at  $4\sigma$ . **d**, Detailed view of the UbiX-product complex in two orientations related by a 90 degree rotation along the horizontal axis. The omit map for the product is shown as green mesh, contoured at  $3.5\sigma$ .



**Figure 3 | Crystal structures of *P. aeruginosa* UbiX<sup>E49Q</sup>.** **a**, Detailed view of the UbiX<sup>E49Q</sup>-FMN-DMAP complex with individual amino acids contributing to the active site structure shown in atom-coloured sticks (carbons colour coded as in Fig. 2a). Two orientations are displayed as in Fig. 2. The omit map for the DMAP substrate is shown as green mesh, contoured at  $4\sigma$ . **b**, Detailed view of the UbiX<sup>E49Q</sup>-FMNH<sub>2</sub>-DMAP complex obtained through rapid flash-cooling following reduction. The omit map for the DMAP substrate is shown as green mesh, contoured at  $3\sigma$ . **c**, Detailed view of the UbiX<sup>E49Q</sup> N5-C1' alkyl adduct species obtained through flash-cooling following reduction. The omit map for the N5-C1' adduct is shown as green mesh, contoured at  $4\sigma$ .

Many of the conformational changes observed in the E49Q N5-dimethylallyl adduct are unlikely to present a significant barrier to the reaction even within the crystals. However, the observed reorientation of the dimethylallyl moiety would require substantial protein breathing motions to occur if achieved through rotation along the dimethylallyl C1-C2 bond. An alternative route through olefin isomerization is possible, via protonation of the C2' atom (adding a *pro-S* hydrogen) with transient formation of a ternary C3' carbocation (Fig. 4, species **III**). The latter could reorient in the absence of active site deformation and be converted to the observed E49Q dimethylallyl-N5 adduct (Fig. 4, species **IVa**) by abstraction of the C2' *pro-R* proton. The substrate-derived phosphate is ideally positioned to establish a proton relay network to the substrate C2', and the ternary C3' carbocation resides within a  $\pi$ -cage made by the FMN dimethylbenzene moiety, Y169 and W200 that could assist carbocation formation.



**Figure 4 | Schematic representation of the proposed UbiX mechanism.** DMAP derived atoms are shown in red. Roman numerals indicate the various intermediate species proposed (see main text). References to individual figures

after roman numerals refer to corresponding crystal structures obtained for the wild type or UbiX<sup>E49Q</sup> mutant.

Surprisingly, despite little difference in structure, crystals of a Y169F variant are severely compromised in their catalytic ability. A similar observation is made when using UbiX<sup>Y169F</sup> variant for *in vitro* reconstitution of Fdc1 activity (Fig. 1e). In the UbiX<sup>Y169F</sup> crystals, an N5–C1' dimethylallyl adduct nearly identical to that seen for the wild-type crystals can be readily obtained, but remains stable for several minutes (Extended Data Fig. 8). This suggests Y169 is key to the conformational changes that follow formation of the first N5 adduct (Fig. 4, species III), possibly through assisting proton transfer via the substrate phosphate to the dimethylallyl C2' or stabilization of the ternary C3' carbocation.

We have been unable to trap any intermediates during formation of the dimethylallyl C3'–flavin C6 bond. We suggest C6 nucleophilic attack on the dimethylallyl C3' carbocation occurs concomitant with or following protonation of the C2' via the bound phosphate (Fig. 4, species IVb). The resulting cyclohexadiene adduct (species V) can form the final product (species VI) through aromatization concomitant with proton abstraction via S15 and E49 (Extended Data Fig. 9). We suggest the E49Q mutation also affects this particular deprotonation, leading to the accumulation of an intermediate (species IVa) preceding this step as observed in the crystals. It is possible that formation of the cyclohexadiene adduct (V) occurs directly following conformational change of the C3' carbocation species (III), via species IVb without intermediate deprotonation/protonation as outlined for species IVa.

In addition to UbiX, a distinct family of flavoenzymes (type 2 isopentenyl diphosphate isomerases) bind isoprene precursors. The latter have been suggested to use the flavin as an acid-base in the interconversion of IPP and DMAPP<sup>15,16</sup>. Instead, our data reveal UbiX has many similarities to terpene synthases, achieving the required isoprenyl chemistry via similar strategies (Fig. 4)<sup>8,17,18</sup>. We suggest formation of an initial allylic carbocation through leaving of the

phosphate group (pyrophosphate in the terpene synthases) is achieved through stabilization of the carbocation species within an active site dominated by  $\pi$ -systems (the  $\pi$ -cage). Furthermore, FMN reduction appears to act as the trigger for phosphate–C1' bond breakage, increasing the electron density of the isoalloxazine ring that stacks with the dimethylallyl substrate. In contrast to the metal-requiring terpene synthases, the leaving group is stabilized by a multitude of ionic interactions. A direct contact with the conserved E140 furthermore suggests phosphate protonation occurs concomitant with C1' allylic carbocation formation and/or N5 nucleophilic attack (Fig. 4, species I<sup>reduced</sup>). Following formation of the N5–C1' bond, UbiX appears to act as a chaperone for substrate reorientation (conversion of species III to IVb), similar to what is postulated to occur for terpene synthases. In UbiX, it appears both the leaving group phosphate and a conserved tyrosine residue (Y169) are involved in catalysing the conformational change of the N5 adduct required for completion of the reaction.

The presence of *ubiX/ubiD*-like genes in the majority of microbes<sup>2,19–21</sup> and the essential role played in prokaryotic ubiquinone biosynthesis<sup>2</sup> suggest an ancient evolutionary origin for the UbiX flavin prenyltransferase chemistry. It remains unclear why dimethylallyl-monophosphate is used by UbiX, as opposed to the universal dimethylallyl-diphosphate<sup>22</sup>. This might serve as a means to regulate product formation via availability of DMAP. Although the isomer isopentenyl monophosphate has recently been shown to occur in certain archaea<sup>23</sup>, the metabolic route to dimethylallyl-monophosphate remains unclear. The unusual biochemical strategy of using a reduced rather than oxidized flavin to act as a nucleophile<sup>9,24,25</sup> ensures formation of a relatively stable N5-alkyl adduct species (prFMN<sup>reduced</sup>), as opposed to the corresponding N5-iminium adduct (prFMN<sup>iminium</sup>). The latter would be prone to hydrolysis before transfer to *apo*-UbiD/Fdc1. Whether non-*ubiD* related enzymes make use of the novel chemical properties of the UbiX product remains to be established.



**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 8 December 2014; accepted 13 May 2015.**

**Published online 17 June 2015.**

1. Lenaz, G. & Genova, M. L. Mobility and function of coenzyme Q (ubiquinone) in the mitochondrial respiratory chain. *Biochim. Biophys. Acta* **1787**, 563–573 (2009).
2. Aussel, L. *et al.* Biosynthesis and physiology of coenzyme Q in bacteria. *Biochim. Biophys. Acta* **1837**, 1004–1011 (2014).
3. Gulmezian, M., Hyman, K. R., Marbois, B. N., Clarke, C. F. & Javor, G. T. The role of UbiX in *Escherichia coli* coenzyme Q biosynthesis. *Arch. Biochem. Biophys.* **467**, 144–153 (2007).
4. Rangarajan, E. S. *et al.* Crystal structure of a dodecameric FMN-dependent UbiX-like decarboxylase (Pad1) from *Escherichia coli* O157:H7. *Protein Sci.* **13**, 3006–3016 (2004).
5. Payne, K. A. P. *et al.* New cofactor supports  $\alpha,\beta$ -unsaturated acid decarboxylation via 1,3-dipolar cycloaddition. *Nature*. <http://www.dx.doi.org/10.1038/nature14560> (2015).
6. Liang, P. H. Reaction kinetics, catalytic mechanisms, conformational changes, and inhibitor design for prenyltransferases. *Biochemistry* **48**, 6562–6570 (2009).
7. Doud, E. H., Perlstein, D. L., Wolpert, M., Cane, D. E. & Walker, S. Two distinct mechanisms for TIM barrel prenyltransferases in bacteria. *J. Am. Chem. Soc.* **133**, 1270–1273 (2011).
8. Gao, Y. & Honzatko, R. B. Peters, terpenoid synthase structures: a so far complete view of complex catalysis. *Nat. Prod. Rep.* **29**, 1153–1175 (2012).
9. Walsh, C. T. & Wenczewicz, T. A. Flavoenzymes: versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.* **30**, 175–200 (2013).
10. Heuts, D. P., Scrutton, N. S., McIntire, W. S. & Fraaije, M. W. What's in a covalent bond? On the role and formation of covalently bound flavin cofactors. *FEBS J.* **276**, 3405–3427 (2009).
11. Mukai, N., Masaki, K., Fujii, T., Kawamukai, M. & Iefuji, H. PAD1 and FDC1 are essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* **109**, 564–569 (2010).
12. Lin, F., Ferguson, K. L., Boyer, D. R., Lin, X. N. & Marsh, E. N. Isofunctional enzymes Pad1 and UbiX catalyse formation of a novel cofactor required by ferulic acid decarboxylase and 4-hydroxy-3-polyphenylbenzoic acid decarboxylase. *ACS Chem. Biol.* **10**, 1137–1144 (2015).
13. Kopec, J., Schnell, R. & Schneider, G. Structure of PA4019, a putative aromatic acid decarboxylase from *Pseudomonas aeruginosa*. *Acta Crystallogr. F* **67**, 1184–1188 (2011).
14. Fraaije, M. W. & Mattevi, A. Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem. Sci.* **25**, 126–132 (2000).
15. Berthelot, K., Estevez, Y., Deffieux, A. & Peruch, F. Isopentenyl diphosphate isomerase: a checkpoint to isoprenoid biosynthesis. *Biochimie* **94**, 1621–1634 (2012).
16. Nagai, T. *et al.* Covalent modification of reduced flavin mononucleotide in type-2 isopentenyl diphosphate isomerase by active-site-directed inhibitors. *Proc. Natl Acad. Sci. USA* **108**, 20461–20466 (2011).
17. Christianson, D. W. Unearthing the roots of the terpenome. *Curr. Opin. Chem. Biol.* **12**, 141–150 (2008).
18. Smanski, M. J., Peterson, R. M., Huang, S. X. & Shen, B. Bacterial diterpene synthases: new opportunities for mechanistic enzymology and engineered biosynthesis. *Curr. Opin. Chem. Biol.* **16**, 132–141 (2012).
19. Lupa, B., Lyon, D., Gibbs, M. D., Reeves, R. A. & Wiegel, J. Distribution of genes encoding the microbial non-oxidative reversible hydroxyarylic acid decarboxylates/phenol carboxylases. *Genomics* **86**, 342–351 (2005).
20. Abu Laban, N., Selesi, D., Rattie, T., Tischler, P. & Meckenstock, R. U. Identification of enzymes involved in anaerobic benzene degradation by a strictly anaerobic iron-reducing enrichment culture. *Environ. Microbiol.* **12**, 2783–2796 (2010).
21. Erb, T. J. Carboxylases in natural and synthetic microbial pathways. *Appl. Environ. Microbiol.* **77**, 8466–8477 (2011).
22. Chang, W. C., Song, H., Liu, H. W. & Liu, P. Current development in isoprenoid precursor biosynthesis and regulation. *Curr. Opin. Chem. Biol.* **17**, 571–579 (2013).
23. VanNice, J. C. *et al.* Identification in *Haloferax volcanii* of phosphomevalonate decarboxylase and isopentenyl phosphate kinase as catalysts of the terminal enzyme reactions in an archaeal alternate mevalonate pathway. *J. Bacteriol.* **196**, 1055–1063 (2014).
24. Sobrado, P. Noncanonical reactions of flavoenzymes. *Int. J. Mol. Sci.* **13**, 14219–14242 (2012).
25. Sun, H. G., Ruszczycky, M. W., Chang, W. C., Thibodeaux, C. J. & Liu, H. W. Nucleophilic participation of reduced flavin coenzyme in mechanism of UDP-galactopyranose mutase. *J. Biol. Chem.* **287**, 4602–4608 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by BBSRC grants (BB/K017802/1 with Shell and BB/M017702/1). We thank Diamond Light Source for access to beamlines (proposal number MX8997) that contributed to the results presented here. S.H. is a BBSRC David Phillips research fellow. N.S.S. is an EPSRC Established Career Fellow and Royal Society Wolfson Award holder. The authors acknowledge the assistance given by IT Services and the use of the Computational Shared Facility and the Protein Structure Facility at The University of Manchester.

**Author Contributions** M.D.W. carried out molecular biology, biophysical and structural biology studies together with K.A.P.P. and D.L. M.D.W. and S.A.M. performed *in vitro* reconstitution experiments. K.F. and S.E.J.R. performed and analysed EPR experiments. S.H. performed DFT calculations. N.J.W.R. and D.K.T. undertook liquid chromatography–mass spectrometry of extracts and with R.G. interpreted the data on substrate–product species. All authors discussed the results with D.P. and N.S.S. and participated in writing the manuscript. D.L. initiated and directed this research.

**Author Information** Coordinates and structure factors have been deposited in the Protein Data Bank under accession numbers 4ZAF, 4ZAV, 4ZAW, 4ZAX, 4ZAG, 4ZAL, 4ZAY, 4ZAN and 4ZAZ. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.L. ([david.leys@manchester.ac.uk](mailto:david.leys@manchester.ac.uk))



## METHODS

**Cloning of *P. aeruginosa* UbiX for heterologous expression in *E. coli*.** The *ubix* gene of *Pseudomonas aeruginosa* was amplified from synthesized DNA (codon optimized for expression in *Escherichia coli*, Genescript) by PCR using Phusion polymerase (NEB). The PCR products were cloned into the NcoI and HindIII restriction sites of the pNIC28a-Bsa4 vector using ligation-independent methods (Infusion HD, Clontech). This generated a construct with an N-terminal polyhistidine tag and TEV protease cleavage site similar to that used in ref 12. The Infusion product was transformed into *E. coli* NEB5 $\alpha$  competent cells (NEB) to allow verification of the expression construct by sequencing (Eurofins). Validated pNIC28-Bsa4-*ubix* was transformed into *E. coli* BL21 (DE3) competent cells (NEB) for protein production.

**Cloning of *A. niger* Fdc1 for heterologous expression in *E. coli*.** The *A. niger* *fdc1* gene was codon optimized to remove codons that were rare in *E. coli* and synthesized (Genscript) in pUC57. The gene was excised out of pUC57 using NdeI and XhoI, cloned into the NdeI and XhoI sites of pET30a using T4 ligase (NEB) and transformed into *E. coli* NEB5 $\alpha$ . Once the sequence of the insert was confirmed, *A. niger* Fdc1 pET30a was transformed into *E. coli* BL21(DE3).

**Mutagenesis of *P. aeruginosa* UbiX for heterologous expression in *E. coli*.** Mutagenesis primers were designed using the QuikChange Primer Design Program (<http://www.genomics.agilent.com/primerDesignProgram.jsp>) and used in conjunction with Phusion-polymerase-mediated PCR. Template DNA was removed by DpnI (NEB) digest and the PCR products were transformed into *E. coli* NEB5 $\alpha$  competent cells to allow mutagenesis to be verified by sequencing. Constructs possessing validated mutations were transformed into *E. coli* BL21 (DE3) competent cells for protein production.

**Heterologous expression of *P. aeruginosa* UbiX (wild-type and variants) and *A. niger* Fdc1.** *E. coli* transformants were grown in 1 l of Luria-Bertani (LB) broth at 37 °C and 180 r.p.m. until an  $A_{600\text{nm}}$  (OD<sub>600</sub>) of 0.6–0.8 was reached. Cultures were induced with 0.3 mM IPTG and grown overnight at 20 °C and 180 r.p.m. Cells were collected by centrifugation at 7,000g for 10 min.

**Purification of *P. aeruginosa* UbiX (wild-type and variants).** Cell pellets were resuspended in 50 mM Tris, 500 mM NaCl pH 8.0 buffer supplemented with lysozyme, RNase, DNase (Sigma) and a Complete EDTA free protease inhibitor cocktail (Roche). The cells were lysed using a French Press cell disrupter at 20,000 p.s.i. followed by centrifugation at 48,000g for 30 min and filtration through a 0.45 micron membrane. The soluble supernatant was loaded on to equilibrated Ni-NTA agarose resin (Qiagen) and washed successively with 4 column volumes of resuspension buffer containing 10 and 40 mM imidazole. The protein was eluted with 250 mM imidazole and analysed by SDS-PAGE to verify sample purity. Imidazole was removed using a 10-DG desalting column (Bio-rad) equilibrated with 20 mM Tris, 200 mM NaCl pH 8.0. Where necessary, the samples were incubated with 10 mM FMN to improve flavin occupancy in the protein. Excess FMN was removed before experimentation by desalting, as described above. The N-terminal polyhistidine tag was retained for all experiments.

**Purification of *A. niger* Fdc1.** *A. niger* Fdc1 was purified as described above using 50 mM Tris, 200 mM NaCl pH 7.5 as the resuspension buffer and 20 mM Tris, 100 mM NaCl pH 7.5 as the desalting buffer.

**UV-visible spectrometry and protein quantification.** UV-visible absorbance spectra were recorded using a Cary UV-visible spectrophotometer. Protein concentrations were estimated from  $A_{280}$  using their respective absorbance coefficients, which were calculated from their primary amino acid sequence using the ProtParam program on the ExPASy proteomics server. *P. aeruginosa* UbiX (wild-type and E49Q)  $\epsilon_{280} = 16,960 \text{ M}^{-1} \text{ cm}^{-1}$ ; Y169F  $\epsilon_{280} = 15,470 \text{ M}^{-1} \text{ cm}^{-1}$  and; W200F  $\epsilon_{280} = 11,460 \text{ M}^{-1} \text{ cm}^{-1}$ . The concentration of *A. niger* Fdc1 and FMN were estimated using  $\epsilon_{280} = 68,870 \text{ M}^{-1} \text{ cm}^{-1}$  and  $\epsilon_{450} = 12,500 \text{ M}^{-1} \text{ cm}^{-1}$ , respectively.

**Analysis of DMAP binding to *P. aeruginosa* UbiX.** Distinct perturbations in the absorbance profile of FMN allowed the binding of DMAP to be monitored using UV-visible spectroscopy. The effects of titrating minimal volumes of 12.5  $\mu\text{M}$  DMAP on the absorbance spectrum of 30  $\mu\text{M}$  UbiX:FMN were recorded and converted to difference spectra by subtracting titrated profiles from the DMAP-free sample. The dissociation constant of DMAP was determined by analysing the total deviation in absorbance between 310 and 540 nm in function of DMAP concentration using a tight binding equation.

**Reduction and oxidation of *P. aeruginosa* UbiX (wild-type and variants) for UV-visible spectroscopy and EPR analysis.** All anaerobic work was performed under  $\text{N}_2$  in a glove box (Belle Technology, UK). 150  $\mu\text{M}$  *P. aeruginosa* UbiX:FMN was rendered anaerobic in 20 mM Tris, 200 mM NaCl pH 8.0 and mixed with 1 mM DMAP to monitor the effect of redox cycling on flavin properties. Sodium dithionite was titrated into the sample to fully reduce the FMN, which was subsequently re-oxidized. UV-visible spectra were recorded at each stage and 300- $\mu\text{l}$  aliquots were isolated for EPR analysis.

**Preparation of cofactor extracts for mass identification.** The oxidized product of UbiX turnover was extracted from *P. aeruginosa* UbiX under aerobic conditions by denaturing the protein with equal volumes of acetonitrile at 70 °C. Precipitate was removed by centrifugation at 16,000g for 10 min. The resulting supernatant was sent for high-resolution mass analysis.

The reduced product of UbiX turnover was generated under anaerobic conditions from a steady-state reaction containing 1 mM FMNH<sub>2</sub>, 1 mM DMAP and 20  $\mu\text{M}$  *P. aeruginosa* UbiX. Protein was removed by precipitation at room temperature as described above. The resulting supernatant was sent for analysis by high-resolution mass spectrometry.

**Stopped-flow kinetics of *P. aeruginosa* UbiX.** The kinetics of a single turnover reaction between UbiX:FMNH<sub>2</sub> and DMAP was studied under anaerobic conditions on a SX18MV-R stopped-flow spectrophotometer (Applied Photophysics, dead time approximately 1 ms) by measuring changes in the flavin absorbance spectrum logarithmically over 60 s using a photodiode array (PDA) detector. 100  $\mu\text{M}$  UbiX:FMN was rendered anaerobic in 20 mM Tris, 200 mM NaCl pH 8.0 and supplemented with 1 mM glucose, 10 U ml<sup>-1</sup> glucose oxidase (Sigma) to guarantee an oxygen-free environment. The protein was reduced with sodium dithionite before being rapidly mixed with equal volumes of 500  $\mu\text{M}$ –4 mM DMAP at room temperature. The spectral intermediates were resolved by singular value decomposition using the Pro-Kineticist program (Applied Photophysics). Data could be fitted with minimum residuals to a three step model ( $A > B > C > D$ ), which allowed the rate constants  $k_{f1}$ ,  $k_{f2}$  and  $k_{f3}$  to be estimated. Species D was identified as a photoinduced adduct and considered an artefact of the prolonged illumination required.

The kinetic parameters of a single turnover reaction between UbiX:prFMN<sup>reduced</sup> and oxygenated buffer were measured using a stopped-flow instrument (see above). Here 100  $\mu\text{M}$  UbiX:FMN was rendered anaerobic (in the absence of glucose oxidase), before being reduced with dithionite and mixed with 1 mM DMAP. The protein was rapidly mixed with equal volumes of 0–100% oxygenated buffer, producing absorbance spectra that could be analysed by single value decomposition. Data were fitted to a one-step model ( $A > B$ ), which allowed the rate constant  $k_{AB}$  to be estimated.

**In vitro reconstitution of *A. niger* Fdc1.** For reconstitution under steady-state conditions, a solution containing 1 mM FMNH<sub>2</sub>, 1 mM DMAP and, respectively, wild-type *P. aeruginosa* UbiX, E49Q or Y169F was incubated with 50  $\mu\text{M}$  *A. niger* apo-Fdc1 in 20 mM Tris, 200 mM NaCl, 1 mM MnCl<sub>2</sub> pH 8.0 under anaerobic conditions. The specific activity of reconstituted Fdc1 with 50  $\mu\text{M}$  cinnamic acid (in 50 mM potassium phosphate, 50 mM KCl pH 6.0; either aerobic or anaerobic) was determined at regular time intervals by measuring the rate of decarboxylation at 270 nm using a Cary UV-visible spectrophotometer.

Reconstitution experiments using a filtered UbiX reaction were carried out using 2 mM FMNH<sub>2</sub> anaerobically incubated overnight with 5 mM DMAP and 50  $\mu\text{M}$  UbiX. This reaction mixture was used to reconstitute *A. niger* apo-Fdc1 (supplemented with Mn<sup>2+</sup>) in a 2:1 molar ratio, assuming complete conversion of FMNH<sub>2</sub> to prFMN<sup>reduced</sup>. To obtain filtrate devoid of UbiX, the reaction mixture was anaerobically filtered in a 10K MWCO centrifugal concentrator (Generon). A negative control with DMAP omitted from the overnight reaction was also performed. Decarboxylase activity was measured using 800  $\mu\text{M}$  aerobic cinnamic acid.

**Crystallization of *P. aeruginosa* UbiX (wild-type and variants).** A total of 7.5 mg ml<sup>-1</sup> *P. aeruginosa* UbiX in 20 mM Tris, 200 mM NaCl pH 8.0 was supplemented with 0.5 mM FMN and mixed with 1 mM DMAP to promote ligand binding in the active site. Initial screening, sitting 0.3  $\mu\text{l}$  protein and 0.3  $\mu\text{l}$  mother liquor next to 50  $\mu\text{l}$  reservoir, yielded a variety of hits after 1–2 days at 25 °C. The best crystals were obtained after 1–2 days in 12% PEG 3350, 150 mM sodium thiocyanate, and 100 mM Tris pH 7.2 at 25 °C. Crystals of the UbiX variants were obtained in the same conditions.

**Diffraction data collection and structure elucidation.** All crystals were cryo-protected in mother liquor supplemented with 10% PEG 200 and flash-cooled in liquid nitrogen. Reaction intermediates were obtained by flash cooling crystals that had been incubated in cryoprotectant containing sodium dithionite. Where periods of long soaking were necessary, the cryoprotectant was also supplemented with FMN and DMAP to reduce dissociation from the crystals. Diffraction data was collected up to 1.4 Å at Diamond beamlines at 100K and processed using the CCP4 suite<sup>25</sup>. Data was reduced and scaled using XDS<sup>26</sup>. The structure of *P. aeruginosa* UbiX was refined using REFMAC5, using 3ZQU as the starting model, and refined by cycles of manual rebuilding in COOT and additional processing in REFMAC5 (ref. 25). Ligand coordinates and definitions were generated using the GlycoBioChem PRODRG2 server (<http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg>). The data and refinement statistics are available in Extended Data Table 1.

**Mass spectrometry.** All solvents were of HPLC-MS grade and purchased from Sigma-Aldrich (Gillingham, UK). HPLC grade formic acid was purchased from Fisher Scientific (Loughborough, UK). Mass spectrometer calibration solution and chromatography columns were purchased from Thermo-Fisher Scientific (Hemel Hempstead, UK).

All UHPLC-MS work was carried out on a Thermo-Finnigan Orbitrap-LTQ XL hybrid mass spectrometer operated in negative ionization mode coupled to a Thermo Accela autosampler (Fisher Scientific, Bremen, Germany). Chromatographic separations were adapted from work carried out by Fu and co-workers<sup>27</sup> and performed on a Thermo Hypersil Gold 2.1  $\mu\text{m}$  C<sub>18</sub> column at a solvent flow-rate of 400  $\mu\text{l min}^{-1}$ . For initial profiling tests the column was eluted with 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The solvent composition during gradient elution was initiated with 5% (B) for 5 min and subsequently ramped to 95% (B) over 15 min, followed by a 5 min isocratic elution at 95% (B) before a return to 95% (A) held for further 5 min for column equilibration. All samples were maintained at 4 °C within the autosampler refrigerator while the column was maintained at 50 °C within the autosampler oven. Mass calibration was carried out in accordance with the manufacturer's guidelines using caffeine, the tetrapeptide MRFA and Ultramark 1621 in an acetonitrile/methanol/acetic acid solution. Acquisition settings for initial profiling were carried out at 60,000 resolution in centroid and ran at 1  $\mu$ -scan per 400 ms in the 100–1,000  $m/z$  range with source gasses set at sheath gas = 40 arbitrary units, aux gas = 5 arbitrary units, sweep gas = 5 (all systems specific arbitrary units). The ESI source voltage was set to 4.2 V, and capillary ion transfer tube temperature set at 275 °C.

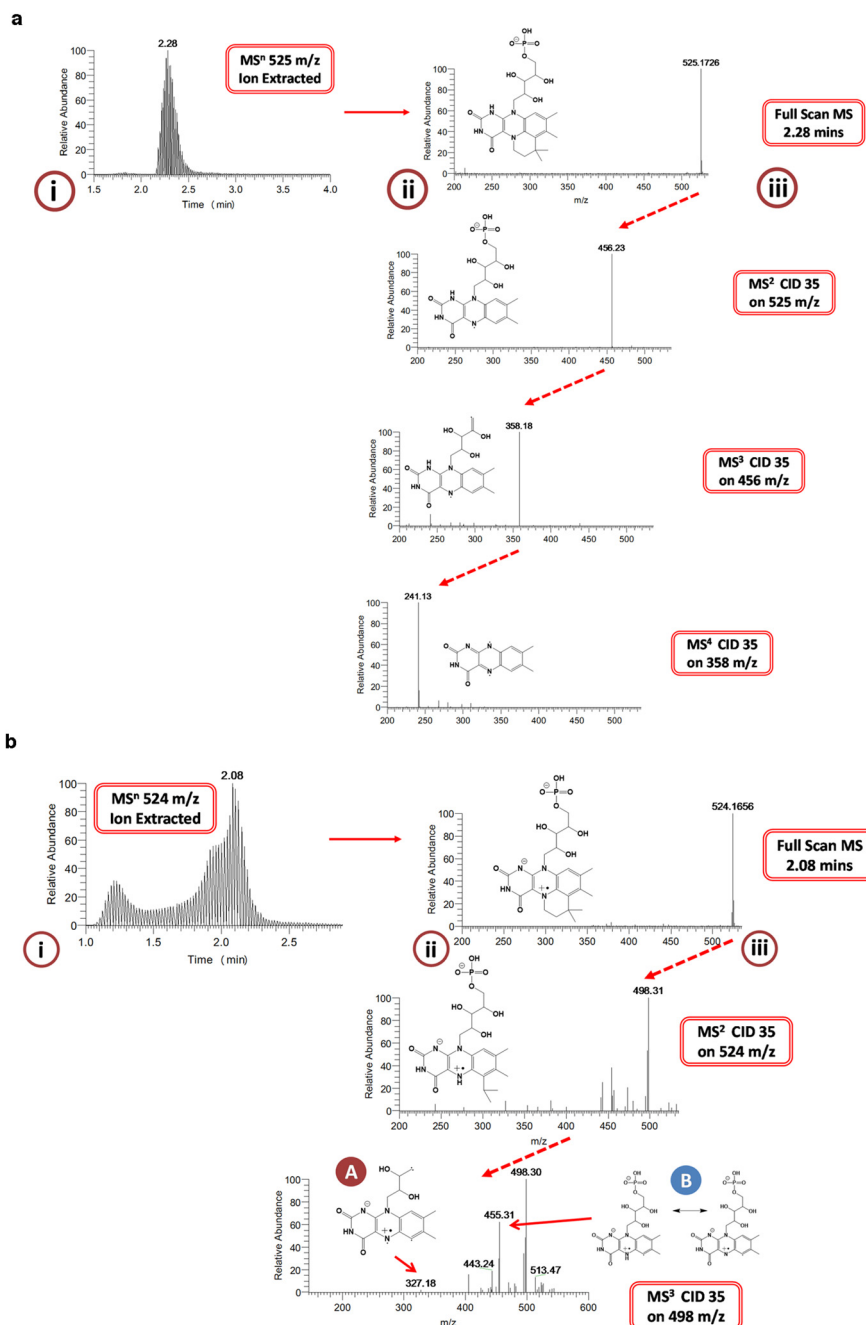
Mass fragmentation analysis was carried out with the same column chemistry, source settings and flow rate but with an isocratic solvent elution of 55% (A)/45% (B), an optimum composition determined for eluting the analyte of interest, based on the initial elution profile described above. MS source, sample storage/column was kept under identical conditions. Collision induced dissociation (CID) settings were set up to trap target ions with an isolation width of 1.0  $m/z$ , normalized collision energy of 35, activation Q of 0.250 and activation time of 30 ms.

**EPR spectroscopy.** Continuous wave X-band (~9.4 GHz) EPR spectra were obtained using a Bruker ELEXSYS E500 EPR spectrometer operating at cryogenic temperatures via an Oxford Instruments ESR900 liquid helium cryostat interfaced with an ITC503 temperature controller from the same manufacturer. Samples were 250  $\mu\text{l}$  in Wilmad 4 mm outer diameter quartz tubes. Sample temperatures were as stated, microwave power was 10  $\mu\text{W}$ , modulation amplitude was 1.5 G. Pulsed electron nuclear double resonance (ENDOR) spectra were obtained at 65 K using a Bruker ELEXSYS E580 spectrometer equipped with an EN 4118X-MD4 dielectric resonator. Temperature control was effected using an Oxford Instruments CF935 cryostat and ITC503 temperature controller. The Davies pulsed ENDOR sequence was employed at intermediate Q using soft microwave pulses and FID detection ( $\pi/2 = 200$  ns) with a 9 ms radiofrequency  $\pi$  pulse.

**DFT calculations.** Density function theory (DFT) models of FMN:DMA adducts were optimized in the gas phase using the (U)BP3LYP/6-311++G(d,p) level of theory implemented in Gaussian 09 (ref. 28). The models consist of an FMN truncated after the 2' carbon and 2-methyl-2-butene, that is, the dephosphorylated DMAP (Extended Data Fig. 4). Structural alignments to the crystal coordinates were performed using Swiss-PdbViewer version 4.1 (ref. 29). Harmonic vibrational frequencies calculated using normal mode analysis were used to confirm that optimized geometries were always in local or global minima.

**Data reporting.** No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

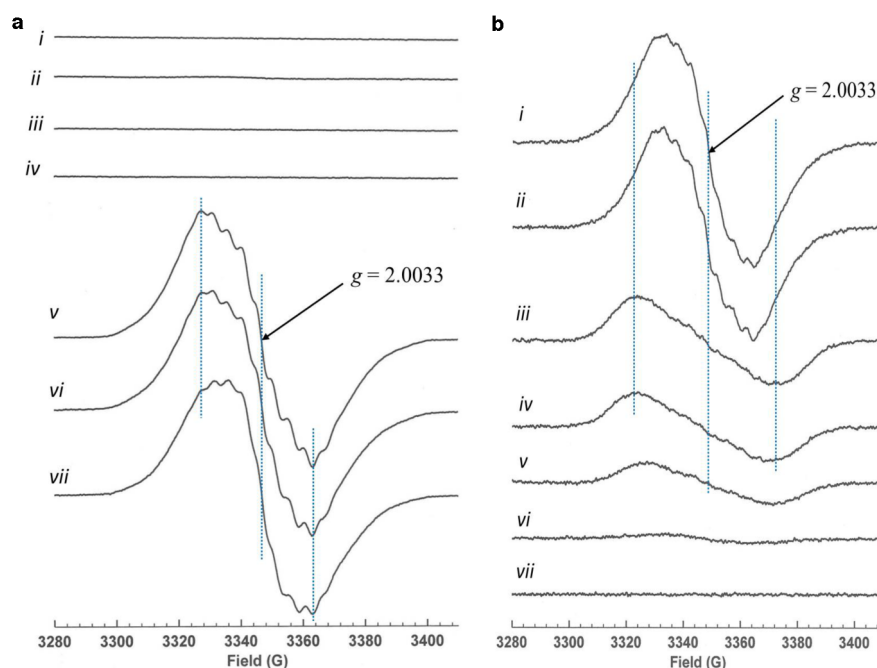
26. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
27. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
28. Fu, G. *et al.* Atomic-resolution structure of an N5 flavin adduct in D-arginine dehydrogenase. *Biochemistry* **50**, 6292–6294 (2011).
29. Frisch, X. M. J. *et al.* Gaussian 09 (Gaussian, Wallingford, Connecticut, revision B.01, 2010).
30. Guex, Y. N. & Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).



### Extended Data Figure 1 | Mass spectrometric analysis of the UbiX product.

**a** Structural elucidation of the reduced UbiX product. From an initial full scan TIC on UbiX extract, a 525  $m/z$  ion extracted chromatogram was created under a gradient elution using H<sub>2</sub>O/acetonitrile both containing 0.1% formic acid indicating a major peak apex at 9.57 min with a 54/46 solvent elution composition (not shown). Subsequent data dependant TIC and 525  $m/z$  scan extracted chromatograms were created under 55% A/45% B isocratic solvent elution and ion extraction between 524.5–525.5  $m/z$  produced a singular peak at 2.28 min displaying an associated full-scan molecular ion peak with  $m/z$  = 525.1726 ( $M^+$  = C<sub>22</sub>H<sub>30</sub>N<sub>4</sub>O<sub>9</sub>P) at a resolution of 58,500 with a mass accuracy of 3.59 p.p.m. Fragmentation of the 525.1726  $m/z$  molecular ion peak in an automated data dependent manner using helium-based chemical-induced dissociation (CID level 35) generated a spectral tree that indicates the removal of the newly formed, more labile, tertiary ring at the MS<sup>2</sup> level. Subsequent removal of the phosphate head group at the MS<sup>3</sup> level was achieved using CID 35 on the 456.23  $m/z$  molecular species with a final MS<sup>4</sup> step using CID 35 on 358.18  $m/z$  completely removing the tail group from the central three-ring

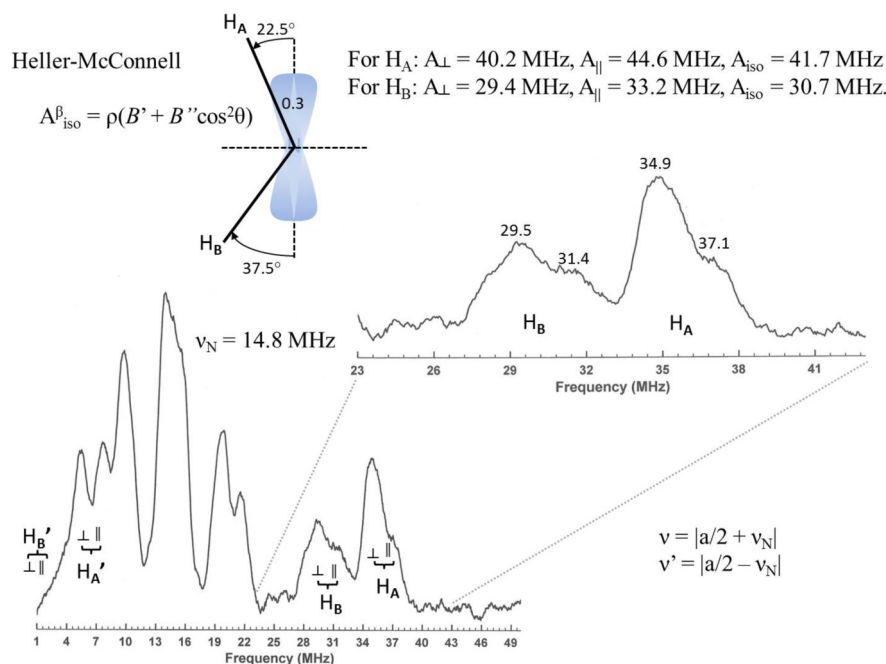
isalloxazine system. **b**, Structural elucidation of the re-oxidized UbiX/Fdc1 cofactor. From an initial full scan TIC on UbiX extract (i), a 524  $m/z$  ion extracted chromatogram was created under a gradient elution using H<sub>2</sub>O/acetonitrile both containing 0.1% formic acid indicating a major peak apex at 9.24 min with a 48/52 solvent composition (not shown). Subsequent data dependant TIC and 524  $m/z$  scan extracted chromatograms (ii) were created under 50% A/50% B isocratic solvent elution and ion extraction between 523.5–524.5  $m/z$  produced a singular peak at 2.08 min displaying an associated full-scan molecular ion peak with  $m/z$  = 524.1656 ( $M^+$  = C<sub>22</sub>H<sub>29</sub>N<sub>4</sub>O<sub>9</sub>P) at a resolution of 58,500 with a mass accuracy of 2.78 p.p.m. Fragmentation of the 524.1656  $m/z$  molecular ion peak in an automated data dependent manner using helium-based chemical-induced dissociation (CID level 35) generated a spectral tree (iii) that indicates the removal of the newly formed, more labile, tertiary ring at the MS<sup>2</sup> level. Subsequent removal of the phosphate head group at the MS<sup>3</sup> level was achieved using CID 35 on the 498.31  $m/z$  molecular species to create 327.18 (A) alongside a sister fragment 455.31 (B) that represents the full removal of the tertiary ring but retaining the phosphate head group.



**Extended Data Figure 2 | EPR spectroscopic analysis of the UbiX radical product.** **a**, X-band continuous wave EPR spectra of UbiX in frozen solution: (i) wild type (WT) as isolated; (ii) WT plus DMAP; (iii) WT reduced with dithionite; (iv) WT + DMAP reduced with dithionite; (v) WT + DMAP reduced with dithionite and reoxidized with oxygen; (vi) Y169F mutant + DMAP reduced with dithionite and reoxidized with oxygen; (vii) W200F mutant + DMAP reduced with dithionite and reoxidized with oxygen. The FMN–DMAP adduct radical is only formed when UbiX is reoxidized in the presence of DMAP and this formation is not affected by mutation of those aromatic residues forming the  $\pi$ -cage that could give rise to Y or W radical species. **b**, X-band continuous wave EPR spectra of frozen solutions of WT

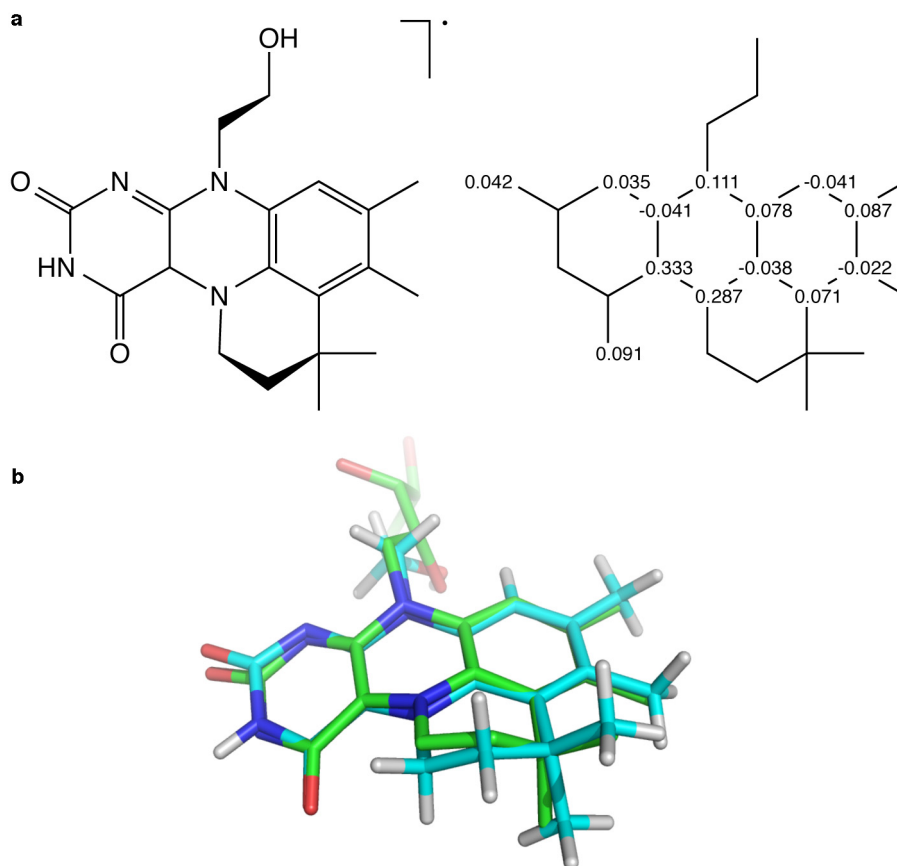
UbiX + DMAP and reduced with dithionite with the addition of potassium ferricyanide to the following concentrations: (i) 260  $\mu$ M; (ii) 160  $\mu$ M; (iii) 50  $\mu$ M; (iv) 40  $\mu$ M; (v) 30  $\mu$ M; (vi) 20  $\mu$ M; (vii) 0  $\mu$ M. Experimental conditions: microwave power 10  $\mu$ W, modulation amplitude 1.5 G, temperature 20 K. Showing the radical can also be formed using chemical oxidation in the absence of oxygen and thus does not arise from a peroxide species generated by the reaction of reduced oxygen species formed when the dithionite sample is exposed to oxygen. An initial radical is formed under these conditions that exhibits a considerably broader EPR signal than the prFMN<sup>radical</sup> and is as yet unidentified.





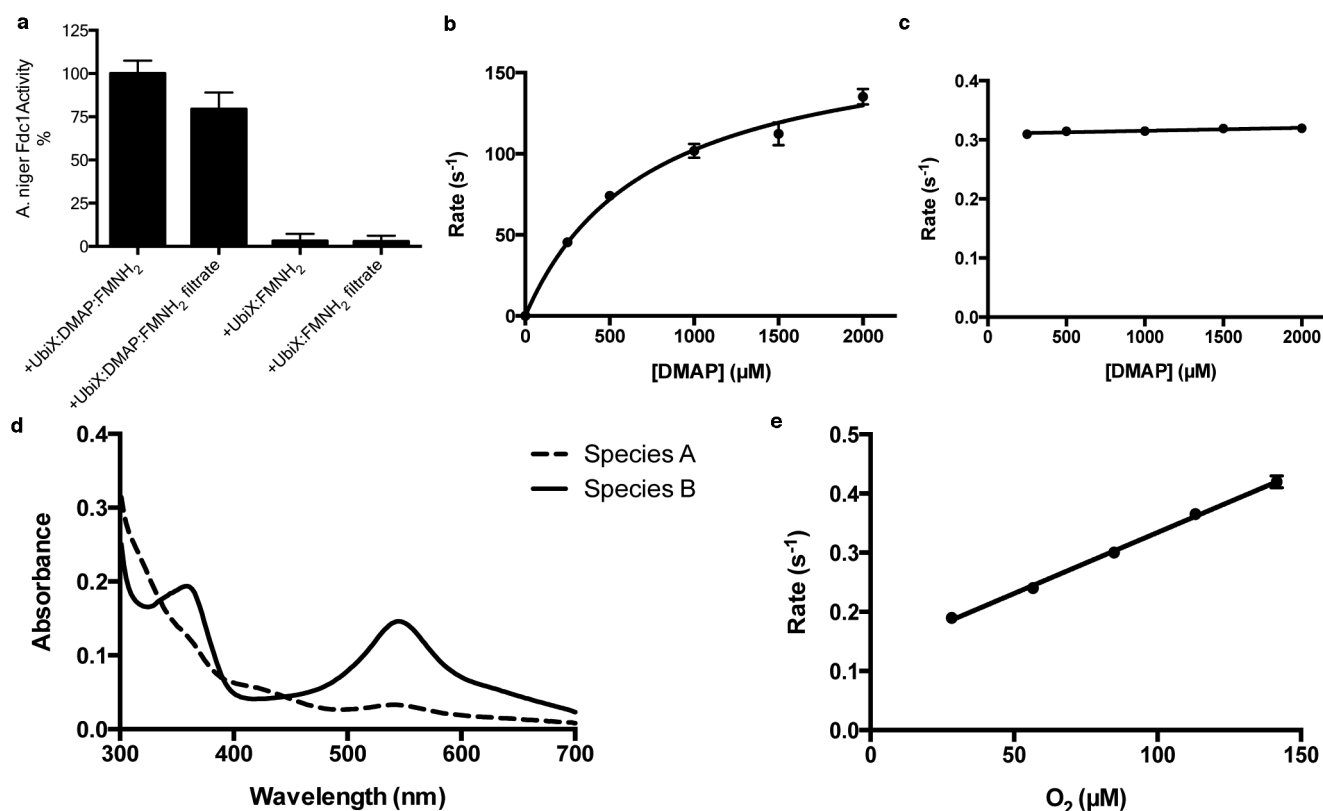
**Extended Data Figure 3 | Pulsed Davies ENDOR spectra of the prFMN<sup>radical</sup>-UbiX complex.** The spectrum was measured at a field equivalent to  $g_{\text{av}} = 2.0033$ . Although a complete assignment of the spectrum requires specific deuteration of FMN and DMAP, the ENDOR spectrum is dominated by two large hyperfine couplings to  $\beta$ -protons indicated as  $H_A$  and  $H_B$ . Using the Heller-McConnell equation the values of the dihedral angles,  $\theta$ , can be determined as shown and are consistent with the orientation of the C1'-protons of the DMAP-derived fragment of the radical observed crystallographically,

as shown in the figure above. The unpaired electron spin density,  $\rho$ , at N5 of the FMN-derived fragment of the radical can also be estimated from the Heller-McConnell equation.  $B'$  is negligible, whereas  $B''$  is thought to have a value of  $\sim 160$ , although studies of  $\beta$ -protons coupled to unpaired electron spin density at a nitrogen atom are rare, giving an unpaired spin density of  $\sim 0.3$ , consistent with calculations and considerably smaller than the unpaired electron spin density of 0.4 or greater expected for C1' of an aromatic amino acid radical.



**Extended Data Figure 4 | DFT modelling of the UbiX product.** Top, DFT model of the purple radical species showing the location of significant atomic spin densities ( $>|0.02|$ ) to the right. The optimized structure (blue carbons) overlaid with the crystal coordinates (green carbons) is shown below. The

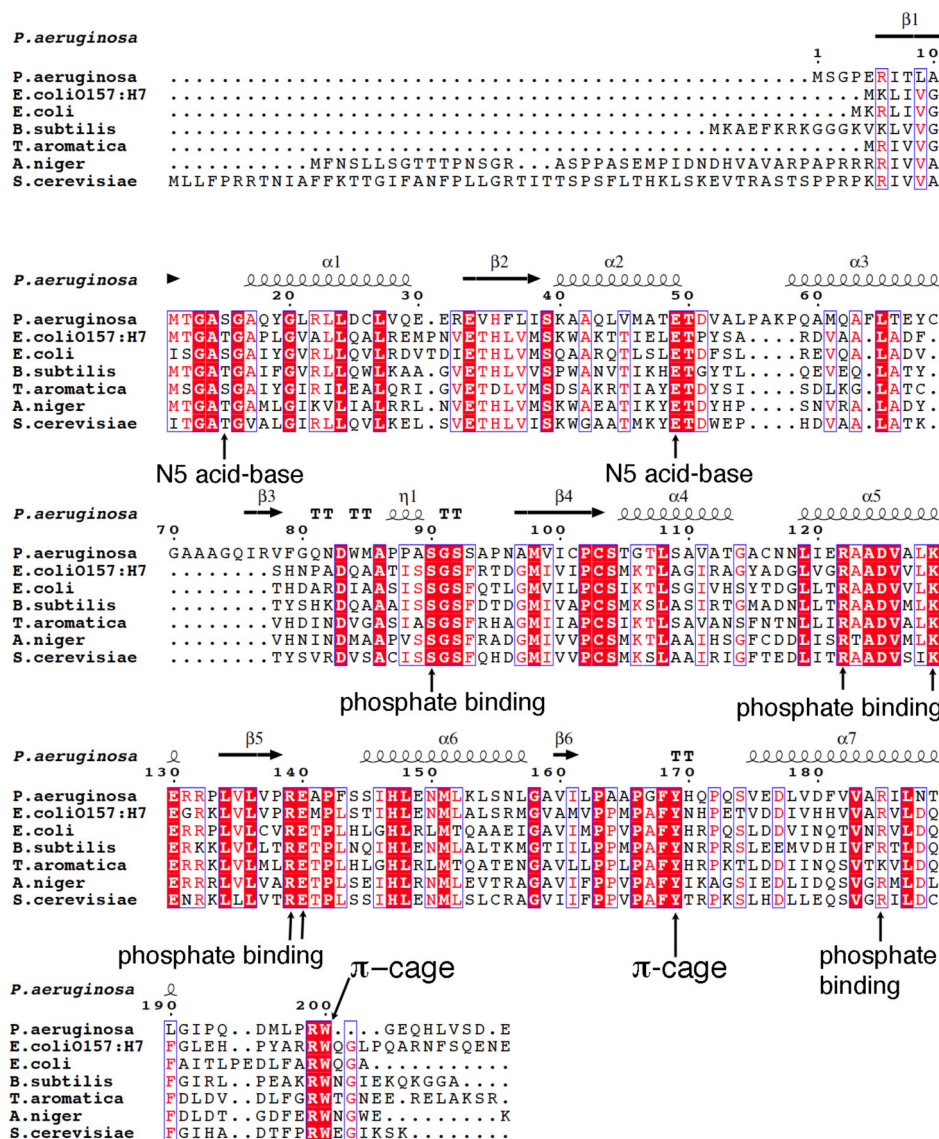
model was geometry optimized in the gas phase using the UB3LYP/6-311++G(d,p) level of theory. Cartesian coordinates of the optimized structure are given in the Supplementary Information.



#### Extended Data Figure 5 | Additional UbiX solution studies.

**a**, Reconstitution of *A. niger* Fdc1 activity with UbiX-prFMN<sup>reduced</sup> and prFMN<sup>reduced</sup> obtained through filtration of a UbiX-prFMN<sup>reduced</sup> reaction. Control reactions are devoid of any DMAP substrate. **b**, Rate of formation of spectral species 2 (see Fig. 1f) in function of DMAP concentration. **c**, Rate of decay of spectral species 2 (see Fig. 1f) in function of DMAP concentration. **d**, Spectral species obtained from singular value decomposition of rapid-scan

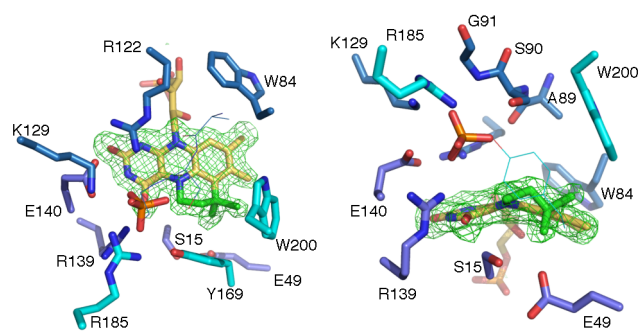
stopped-flow spectrophotometric data following mixing of UbiX-prFMN<sup>reduced</sup> with oxygenated buffer. **e**, The rate of purple radical (species B in panel **d** of this figure) formation as obtained from singular value decomposition of rapid-scan stopped-flow spectrophotometric data following mixing of UbiX-prFMN<sup>reduced</sup> with oxygenated buffer has a linear dependence on oxygen concentration. Error bars are s.e.m.,  $n = 3$ .



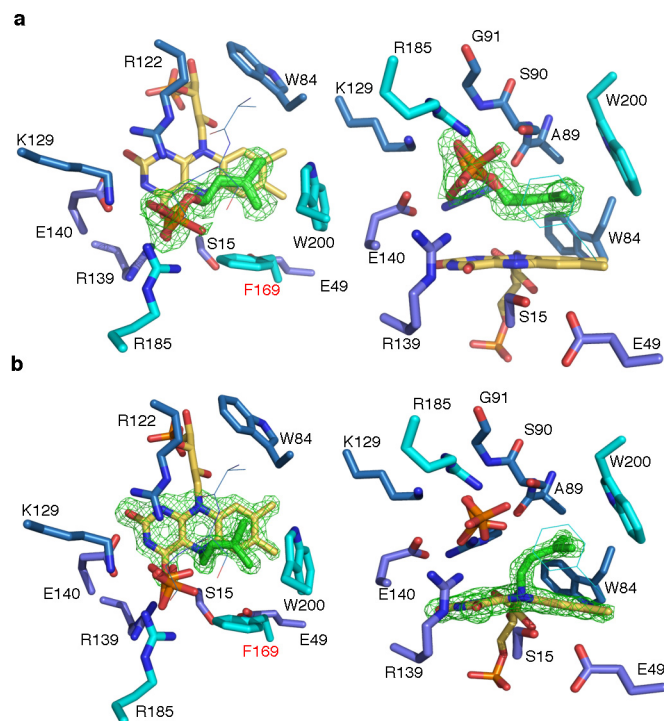
**Extended Data Figure 6 | Multiple sequence alignment of UbiX/Pad1 enzymes from selected bacterial or fungal species.** *Pseudomonas aeruginosa* UbiX (NP\_252708), *Escherichia coli* O157:H7 EcdB (NP\_311620), *Escherichia coli* UbiX (YP\_490553), *Bacillus subtilis* BsdB (WP\_009966530), *Saccharomyces cerevisiae* Pad1 (AAB64980), *Aspergillus niger* PadA1 (ABN13117), and orf8 from the *Thauera aromatica* phenylphosphate

carboxylase gene cluster (PAAD\_THAAR). Conserved residues involved in phosphate binding, N5 polar network or formation of the substrate binding  $\pi$ -cage are indicated by labelled arrows. Secondary structure elements of *P. aeruginosa* UbiX crystal structure are shown. Alpha helices and  $3_{10}$  helices (denoted as  $\eta$ ) are shown as squiggles,  $\beta$ -strands as arrows and  $\beta$ -turns as TT.





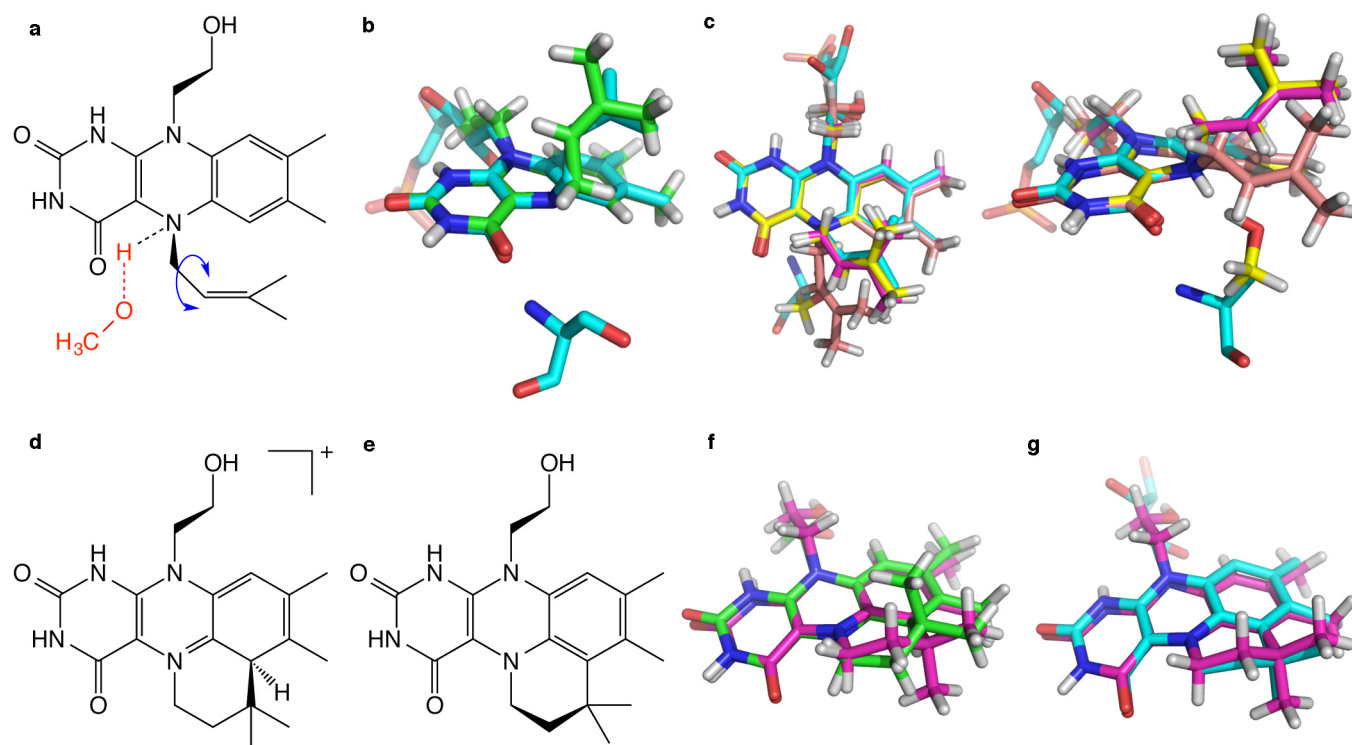
**Extended Data Figure 7 |** Crystal structure of *P. aeruginosa* UbiX-FMN-DMAP flash cooled to 100 K at 30 s following complete reduction by sodium dithionite. Two orientations are displayed as in Fig. 2. The omit map for the prFMN<sup>reduced</sup> product is shown as a green mesh, contoured at  $4\sigma$ .



**Extended Data Figure 8 | Crystal structures of *P. aeruginosa* UbiX<sup>Y169F</sup>.**

**a.** Detailed view of the UbiX<sup>Y169F</sup>-FMN-DMAP complex with individual amino acids contributing to active site structure shown in atom-coloured sticks (carbons colour coded as in Fig. 2a). Two orientations are displayed as in Fig. 2.

The omit map for the DMAP substrate is shown as green mesh, contoured at 4σ. **b.** Detailed view of the UbiX<sup>Y169F</sup> N5-C1' adduct species obtained through flash-cooling following reduction. The omit map for the N5-C1' adduct is shown as a green mesh, contoured at 4σ.



**Extended Data Figure 9 | DFT models of proposed intermediate species in the UbiX reaction.** **a**, DFT models of species **II** and **IVa** (as defined in Fig. 4). Conversion from **II** to **IVa** is achieved by  $\sim 180^\circ$  rotation about  $C1'-C2'$  (blue arrow) and the  $N5-H$  and methanol species (red) are only found in species **IVa** models. **b**, Overlay of the species **II** DFT model (green carbons) with the crystal coordinates of species **II** and **S15** (teal carbons). **c**, Three DFT models of **IVa** were examined and two orthogonal projections are shown overlaid with the crystal coordinates (teal carbons): (**Vi**, yellow carbons) with a methanol analogue of **S15** (a, in red) with the  $C-N5$  distance fixed to the crystallographic distance of 4.0 Å; (**Vii**, magenta carbons) with  $N5$  protonated (no methanol), and (**Viii**, light pink carbons) with  $N5$  deprotonated and no methanol. DFT

model of species **V** and **VI** are shown in **d** and **e**, respectively, and are overlaid in **f** (**V** green carbons, **VI** magenta carbons). **g**, Overlay of the species **VI** DFT model (magenta carbons) with the crystal coordinates (teal carbons). Models were geometry optimized in the gas phase using the B3LYP/6-311++G(d,p) level of theory. Harmonic vibrational frequencies calculated using normal mode analysis were used to confirm that optimized geometries of all species were in local or global minima. In the case of species **Vi**, 'ModRedundant' optimisation was performed to fix the  $C-N5$  distance and one imaginary frequency of  $67.60\text{ cm}^{-1}$  was observed. Cartesian coordinates of the optimized structures are given in the Supplementary Information.

Extended Data Table 1 | Data collection and refinement statistics

| UbiX  | WT<br>FMN:DMAP            | WT<br>N5 adduct 5 s<br>following<br>reduction | WT<br>N5/C6<br>adduct, 30 s<br>following<br>reduction | WT<br>N5/C6<br>adduct, co-<br>crystallised<br>(radical) | E49Q<br>FMN:DMAP          | E49Q<br>FMNH <sub>2</sub> :DM<br>AP, < 5 s<br>following<br>reduction | E49Q N5<br>adduct >15 s<br>following<br>reduction | Y169F<br>FMN:DMAP         | Y169F<br>N5 adduct<br>>15 s<br>following<br>reduction |
|---|---------------------------|---|---|---|---------------------------|--|---|---------------------------|---|
| PDB code  | 4ZAF                      | 4ZAV  | 4ZAW  | 4ZAX  | 4ZAG                      | 4ZAL   | 4ZAY  | 4ZAN                      | 4ZAZ  |
| <b>Data collection</b>                              |                           |   |   |   |                           |  |   |                           |   |
| Space group   | F 2 3                     | F 2 3   | F 2 3   | F 2 3   | F 2 3                     | F 2 3  | F 2 3   | F 2 3                     | F 2 3   |
| Cell dimensions                                     |                           |   |   |   |                           |  |   |                           |   |
| <i>a</i> , <i>b</i> , <i>c</i> (Å)                  | 141.9                     | 142.18  | 142.74  | 141.79  | 142.31                    | 142.07   | 142.02  | 141.73                    | 142.26  |
| $\alpha$ , $\beta$ , $\gamma$ (°)                   | 90                        | 90  | 90  | 90  | 90                        | 90   | 90  | 90                        | 90  |
| Resolution (Å)                                      | 70.95-1.71<br>(1.75-1.71) | 50.27-1.40<br>(1.44-1.40)                     | 43.04-1.89<br>(1.94-1.89)                             | 42.75-1.61<br>(1.65-1.61)                               | 32.65-1.68<br>(1.72-1.68) | 42.84-1.62<br>(1.66-1.62)  | 32.58-1.54<br>(1.58-1.54)                         | 70.86-1.76<br>(1.81-1.76) | 50.30-1.45<br>(1.49-1.45)                             |
| <i>R</i> <sub>pim</sub>                             | 2.7 (29.8)                | 2.7 (29.7)                                    | 4.3 (33.0)  | 2.4 (33.5)  | 3.5 (35.6)                | 2.4 (29.2)   | 2.1 (26.6)  | 2.6 (36.0)                | 2.8 (35.6)  |
| <i>I</i> / $\sigma$ <i>I</i>                        | 18.7 (2.8)                | 16 (2.7)                                      | 13.4 (2.7)  | 20.1 (2.5)  | 13.5 (2.2)                | 16.9 (2.8)   | 20.3 (3.0)  | 19.3 (2.4)                | 15.7 (2.4)  |
| Completeness (%)                                    | 100 (100)                 | 99.6 (100)                                    | 99.9 (99.9)   | 99.9 (100)  | 99.9 (99.9)               | 99.8 (100)   | 100 (100)   | 99.9 (100)                | 99.9 (99.9)   |
| Redundancy  | 6.7 (6.8)                 | 6.8 (6.8)                                     | 6.7 (6.7)   | 5.6 (5.4)   | 5.6 (5.5)                 | 6.6 (6.7)  | 6.7 (6.7)   | 6.7 (6.4)                 | 5.5 (5.4)   |
| <b>Refinement</b>                                   |                           |   |   |   |                           |  |   |                           |   |
| Resolution (Å)                                      | 70.95-1.71<br>(1.75-1.71) | 50.27-1.40<br>(1.44-1.40)                     | 43.04-1.89<br>(1.94-1.89)                             | 42.75-1.61<br>(1.65-1.61)                               | 32.65-1.68<br>(1.72-1.68) | 42.84-1.62<br>(1.66-1.62)  | 32.58-1.54<br>(1.58-1.54)                         | 70.86-1.76<br>(1.81-1.76) | 50.30-1.45<br>(1.49-1.45)                             |
| No. reflections                                     | 24314<br>(1295)           | 44084<br>(2399)                               | 18350<br>(990)  | 29027<br>(1541)   | 25816<br>(1402)           | 28593<br>(1525)  | 33304<br>(1756)                                   | 22355<br>(1074)           | 40068<br>(2066)                                       |
| <i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> | 13.74/17.23               | 9.77/12.90                                    | 14.76/19.16   | 14.36/17.10   | 15.15/17.97               | 13.97/16.34  | 14.54/17.60                                       | 14.24/18.20               | 10.26/14.72   |
| No. atoms   |                           |   |   |   |                           |  |   |                           |   |
| Protein   | 1621                      | 1627  | 1554  | 1620  | 1641                      | 1664   | 1727  | 1593                      | 1658  |
| Ligand  | 41                        | 36  | 36  | 36  | 41                        | 41   | 36  | 41                        | 36  |
| Ion   | 13                        | 18  | 13  | 14  | 4                         | 6  | 15  | 7                         | 20  |
| Water   | 139                       | 194   | 100   | 127   | 141                       | 139  | 144   | 116                       | 156   |
| B-factors   |                           |   |   |   |                           |  |   |                           |   |
| Protein   | 23.22                     | 17.562  | 24.13   | 22.685  | 23.492                    | 26.251   | 22.928  | 26.577                    | 20.892  |
| Ligand  | 20.21                     | 14.938  | 19.286  | 26.107  | 21.486                    | 23.631   | 19.212  | 21.411                    | 17.141  |
| Ion   | 38.34                     | 29.653  | 30.154  | 25.756  | 25.435                    | 60.387   | 28.885  | 34.894                    | 28.034  |
| Water   | 33.289                    | 34.115  | 29.071  | 31.840  | 32.838                    | 36.643   | 32.833  | 37.493                    | 32.705  |
| R.m.s deviations                                    |                           |   |   |   |                           |  |   |                           |   |
| Bond lengths (Å)                                    | 0.0253                    | 0.0258  | 0.0209  | 0.0260  | 0.0229                    | 0.0262   | 0.0262  | 0.0222                    | 0.0250  |
| Bond angles (°)                                     | 2.1189                    | 1.9826  | 1.9719  | 2.4312  | 2.1375                    | 2.3693   | 2.3773  | 2.4068                    | 2.2280  |





► military, academia, hospital-based research organizations and health-care-policy think tanks. Many nurse scientists work at university nursing or medical schools, or at large clinical-research centres that are affiliated with them.

Outside these sectors, some nurse scientists work for pharmaceutical and medical-device companies to help to run clinical trials. Others head up scientific programmes at government organizations, including drug regulators and infectious-disease agencies or national health administrations. Health-management and insurance companies, as well as non-profit advocacy groups such as the American Heart Association, also seek nurse scientists to oversee their research programmes. In academia, nurse scientists are in high demand as lecturers, thanks to globally low numbers of nursing faculty members who are needed to train the next generation.

Nurse scientists pursue research that spans molecular biology, physiology, medical imaging, and public health and policy. Nurses who want to engage in research will need to identify a relevant PhD programme and determine how best to balance nursing practice with their studies. Postdoctoral-research experience is not required for many of the nurse-scientist careers listed above, but it is necessary for nurse scientists who want to run their own research groups at research-intensive universities.

Some nurse scientists find it difficult to make time to continue seeing patients in the clinic, whereas others manage to make it work synergistically. Rapid translation and adoption of nursing-research findings can quickly change bedside practice and bring nurse scientists immense gratification. With various career options from which to choose, the common thread that bonds nurse scientists is a focus on studies that improve patient care and, ultimately, quality of life.

## TESTS TO TREATMENTS

After they receive a PhD, some nurse scientists give up seeing patients to run full-time research programmes. But in many cases, their research studies or administrative duties keep them in close contact with clinical practice. As a senior lecturer at Queensland University of Technology in Brisbane, Australia, Kimberly Alexander oversees the nursing school's cancer-nursing curriculum, while also investigating how genetics influence the symptoms and outcomes of cancer treatments. Similarly, Anna Axelin, a neonatal-care nurse at the University of Turku in Finland, no longer works at the bedside, yet her research on providing intensive care for premature babies keeps her close to patients and their families, and on the hospital ward weekly.

Patricia Grady, a neuroscientist and director of the NINR (see "The price of quality of life"), says that nurse scientists who wish to remain in the clinic must find ways to efficiently blend a research question with a well-matched

## FUNDS FOR NURSES

### *The price of quality of life*

One of the least known of the 27 US National Institutes of Health (NIH) centres, the National Institute of Nursing Research (NINR) operates on an annual budget of about US\$144 million. That budget has nearly tripled in the past 20 years, but it is a tiny slice compared with the almost \$5 billion that the NIH gives to the National Cancer Institute, for example. About 10% of NINR funds support a small internal programme of seven research groups, and the rest of the budget is given in the form of grants to fund external research and training programmes.

Unlike the many other institutes, however, the NINR is not focused on a particular disease area, life stage or organ system. "Our science is primarily

clinical research, but we are what can be described as disease-agnostic," says Patricia Grady, director of the NINR in Bethesda, Maryland. "It's all about the science of what causes symptoms, how can you intervene and how can you help people live with chronic symptoms."

NINR-funded research has uncovered differences in how men and women respond to pain medicine; developed better tools for end-of-life and palliative care; and led to programmes that help adolescents to better manage their diabetes.

Grady says that both the evolution of health-care delivery and the ageing of populations worldwide ensures nursing researchers a bright future. **K.P.**

patient population. US Navy commander Jason McGuire, a certified nurse anaesthetist who is stationed at the Walter Reed National Military Medical Center in Bethesda, Maryland, has managed to do just that. He chose to research a condition that affects veterans who are undergoing general anaesthesia for surgery.

In McGuire's work with combat veterans, he noticed a clinical phenomenon that was missing from the scientific literature: his patients experienced 'emergence delirium' at a much higher rate than did the general population. With this condition, veterans wake up from anaesthesia in a confused and combative state. It can lead to physical and cognitive complications, both during recovery and later in life. "I wanted to find out what the incidence was in our military population and how we can decrease the chance of it happening," says McGuire, chief of nursing science and clinical inquiry at the Walter Reed centre.

**Rapid translation and adoption of nursing-research findings can quickly change bedside practice.**

In 2008, the Navy considered McGuire's question important enough to send him to get a PhD at the University of San Diego, California. Gaps in scientific knowledge of nursing practices are common, and often become dissertation topics for nurse scientists. Like many nurses, McGuire chose to pursue the applied side of his research question. "Am I really going to figure out why this delirium occurs? Or can I move a little quicker, to something that will decrease the incidence?" he asked.

Now, four years after receiving his PhD, he is beginning a randomized controlled trial to test whether giving a specific sedation

medication during surgery — before patients enter recovery — can reduce the risk of emergence delirium. By the end of the 3-year, 370-patient study, McGuire might have devised a new standard of surgical care for combat veterans.

Such immediate application of research findings to clinical practice is particular to nursing science and contrasts sharply with related fields, such as medicine, dentistry or veterinary science, in which translation of results into clinical interventions and treatments can take decades. This rapid application makes nursing research especially attractive to researchers who want to see their work improve lives right away.

Quick adoption comes largely through the close ties that many nurse scientists keep with their colleagues who work at the bedside. Of the nearly three million registered nurses in the United States, Grady estimates that only 1% are nurse scientists. In 2013, the US Health Resources and Services Administration reported that almost two-thirds of nursing schools had restricted student enrolments because of faculty shortages (see [go.nature.com/fwcli1](http://go.nature.com/fwcli1)). "We need more of a critical mass," she says.

It is clear that increasing the number of nurse researchers would help to improve patient care. In the decade that Axelin spent in neonatal intensive-care units (NICUs), she noticed a poignant, recurring pattern: during painful procedures such as an intravenous insertion, mothers would weep outside in the hallway while their babies cried inside the hospital room. "It's common practice to ask parents to step outside. But they're both crying," she says. "I thought, 'There must be some solution.'"

Her dissertation research found that parents who held their premature babies during medical procedures alleviated the



children's pain as effectively as did the standard long-accepted protocol — giving babies a sugar solution before a procedure. Within a year of that discovery, almost every NICU in Finland was using Axelin's parental-hold technique, and she had received calls from hospitals in Sweden, Switzerland, Norway, Estonia and Canada that all wanted to implement it.

"It's important for nurses to know and notice that they can make a change in care," says Axelin. "In nursing research, it's all about a holistic view of the patients we work with every day."

Alexander echoes that sentiment, emphasizing that nursing science is not about discovering new therapies or drugs to treat a particular disease. Rather, it focuses on understanding people's responses to treatments and alleviating symptoms to improve quality of life. Her research aims to find ways to tailor nursing care to each patient who has cancer.

Grady says that the NINR's vision for nursing science is to generate a knowledge base that will improve clinical patient care and help to shape US health-care policy. Although many nursing interventions might seem obvious in hindsight — such as parents holding their babies in the NICU — these care procedures must still be tested under rigorous conditions to prove that they lead to better patient outcomes than do current standards of care.

In some ways, nursing science is playing catch-up as researchers race to fill in holes left in the scientific literature — for instance, how transport by helicopter might alter a trauma-patient's physiology. The field is also

poised to determine how nursing care might be delivered to patients now — and in the future — through smartphones and other forms of digital technology.

### IN THE THICK OF IT

The US military has been a front-runner in training nurse scientists, with the US Army, Navy and Air Force each sending one or two nurses for doctoral work each year for the past decade. Although military nurse scientists run research programmes that are designed to address military medicine, situations and populations, their work spans a wide swathe of disciplines, from physiology to public health. Navy Commander Virginia Blackman, a crit-

**"In nursing research, it's all about a holistic view of the patients we work with every day."**

ical-care nurse scientist at the Walter Reed centre, studies how different pain-management protocols used on trauma patients in the field might

affect their long-term development of chronic pain or post-traumatic stress disorder. Like many nurse scientists in hospitals, Blackman advises staff nurses on how best to convert the latest research findings into practice.

As a public-health nurse scientist in the US Air Force, Lieutenant Colonel Jennifer Hatzfeld has worked on projects as diverse as boosting mammogram rates in military wives and — while deployed to Kandahar, Afghanistan — designing protocols for treating blast-fragment wounds. Now based at Fort Detrick, Maryland, she manages the Department of Defense's 'en route' care research portfolio, which focuses on improving care for patients during air or ground transport.

"Some days, it's a little bit like managing somebody's cheque book," Hatzfeld says of her job as a programme manager: she decides which projects to fund. "But I love it because it's an opportunity to strategically guide the future of en route care."

She says that nursing usually attracts those who want to take care of others and contribute to making the world a better place — not necessarily those who see themselves pursuing a doctoral degree. But nurses have important roles in research, she says. They bring unique perspectives to research teams because they are usually closest to the final stages of a study, when an intervention is delivered to a patient or when carers are helping someone to cope with side effects. "There are questions that we need asked and answered that are specific to nursing," Hatzfeld says. "If we don't do it, then we are allowing physicians or other disciplines to ask those questions with what might not be the right perspective or approach." ■

**Kendall Powell** is a freelance writer based in Lafayette, Colorado.

## TRADE TALK

### Career doctor



*After finishing a PhD and a postdoctoral fellowship in virology, Thomas Magaldi forged a career helping science graduate students and postdocs to plan their paths. He is now the career-services administrator at the Memorial Sloan Kettering Cancer Center in New York City.*

#### Why did you go to graduate school?

I wanted to be a professor at a small liberal-arts college. I learned that I needed a PhD to do that.

#### How did that work out?

During my PhD programme, I started to question my goals. I helped to develop a career-networking group for science trainees at Yale University in New Haven, Connecticut, and I explored other career options, including a science-policy fellowship. But every fellow I spoke to had done something interesting outside the laboratory; I knew that I wouldn't be competitive.

#### So what did you do?

While finishing my thesis, I decided that I should also pursue skills that would set me apart. I did an internship for the US Department of State on how the United States could help Mongolia to create a viable science-education programme. I spent my nights doing that during the last few months of my PhD. Then, for my postdoc, I chose a mentor who would keep me excited about science and who was in Washington DC, where I would have access to policymakers. I volunteered to meet with my local congressman about raising the US science budget, and I taught as an adjunct professor.

#### How did all of that experience help you?

When my wife and I had a baby, we needed more than my postdoc salary and we wanted to be near family. I realized that I had already built a competitive CV for many positions outside the lab, but still requiring a science PhD and postdoc training. The specific scientific work I had done would not distinguish me, but my other experiences would. For the jobs that I wanted, I was ready. ■

#### INTERVIEW BY MONYA BAKER

This interview has been edited for length and clarity; see [go.nature.com/5xbdz6](http://go.nature.com/5xbdz6) for more.

SHOJA LAK



Anna Axelin at the University of Turku, Finland.

## ADDICTION

EXIT ↑

Produced with support from:



Breaking the bonds  
of dependency



# natureOUTLOOK

## ADDICTION

25 June 2015 / Vol 522 / Issue No 7557



Cover art: Nik Spencer

### Editorial

Herb Brody,  
Michelle Grayson,  
Kathryn Miller,  
Brian Owens

### Art & Design

Wesley Fernandes,  
Mohamed Ashour,  
Andrea Duffy

### Production

Karl Smart,  
Ian Pope

### Sponsorship

Stephen Brown,  
David Bagshaw,  
Samantha Morley

### Marketing

Hannah Phipps

### Project Manager

Anastasia Panoutsou

### Art Director

Kelly Buckheit Krause

### Publisher

Richard Hughes

### Chief Magazine Editor

Rosie Mestel

### Editor-in-Chief

Philip Campbell

Addiction can devastate the lives of people and their families. Researchers are disentangling its myriad causes and developing new treatments.

Addiction tends to run in families, but scientists are finding that there is no simple 'addictive personality'. Instead, factors that include genes, character traits and early life experiences combine to make the inheritance of addiction a complicated problem (see page S48).

As an addiction develops it changes the brain: neural circuits related to pleasure and reward are hijacked and rewired (S46). Researchers are studying people from birth to try to tease out how these changes affect, and are affected by, brain development — and how they might be reversed (S50).

Treatments for addiction are becoming more sophisticated, but still face major challenges to acceptance. Medication can help to wean people from their addiction (S53). One controversial but effective technique is to reward people for staying clean (S57). Just as important as repairing the addicted brain is fixing the social environment in which people susceptible to addiction live (S56). Another approach is to make the drugs themselves harder to misuse (S60).

It is not only substances that can be addictive. Gambling is, so far, the only behaviour that has been recognized as an addiction, but researchers are considering adding Internet use, sex and shopping to that list (S62). But despite all the progress, the questions still to be answered are daunting (S63).

We are pleased to acknowledge the financial support of the US National Institute on Alcohol Abuse and Alcoholism, National Institute on Drug Abuse, and Reckitt Benckiser Pharmaceuticals Inc., a subsidiary of Indivior PLC, in producing this Outlook. As always, *Nature* retains sole responsibility for all editorial content.

**Brian Owens**

*Contributing editor*

## CONTENTS

### S46 AETIOLOGY

#### The hijacked brain

The impact of substances and behaviours on neural networks

### S48 GENETICS

#### No more addictive personality

The inheritance of addiction is complex

### S50 NEUROSCIENCE

#### Rewiring the brain

The changes caused by addiction

### S53 PHARMACOTHERAPY

#### Quest for the quitting pill

Medication that could curb cravings

### S56 PERSPECTIVE

#### Beyond the neural circuits

Psychosocial skills are as important as medicine, writes Kenneth E. Leonard

### S57 CONTINGENCY MANAGEMENT

#### Why it pays to quit

Cash can help people to stay drug-free

### S60 TECHNOLOGY

#### Barriers to misuse

The development of tamper-proof pills

### S62 PERSPECTIVE

#### Behavioural addictions matter

Compulsive habits deserve dedicated research, says Marc Potenza

### S63 RESEARCH CHALLENGES

#### 4 big questions

The main themes under investigation

## RELATED ARTICLES

Selected papers on addiction published in *Nature*-affiliated journals.

### S64 The development and maintenance of drug addiction

*R. A. Wise & G. F. Koob*

### S73 VTA CRF neurons mediate the aversive effects of nicotine withdrawal and promote intake escalation

*T. E. Grieder et al.*

### S81 Gray-matter volume, midbrain dopamine D2/D3 receptors and drug craving in methamphetamine users

*A. M. Morales et al.*

*Nature Outlooks* are sponsored supplements that aim to stimulate interest and debate around a subject of interest to the sponsor, while satisfying the editorial values of *Nature* and our readers' expectations. The boundaries of sponsor involvement are clearly delineated in the *Nature Outlook* Editorial guidelines available at [go.nature.com/e4dwzw](http://go.nature.com/e4dwzw)

#### CITING THE OUTLOOK

Cite as a supplement to *Nature*, for example, *Nature* Vol. XXX, No. XXXX Suppl., Sxx–Sxx (2015).

#### VISIT THE OUTLOOK ONLINE

The *Nature Outlook Addiction* supplement can be found at <http://www.nature.com/nature/outlook/addiction>. It features all newly commissioned content as well as a selection of relevant previously published material.

All featured articles will be freely available for 6 months.

#### SUBSCRIPTIONS AND CUSTOMER SERVICES

For UK/Europe: Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants, RG21 6XS, UK. Tel: +44 (0) 1256 329242. Subscriptions and customer services for Americas — including Canada, Latin America and the Caribbean: Nature Publishing Group, 75 Varick St, 9th floor, New York, NY 10013-1917, USA. Tel: +1 866 363 7860 (US/Canada) or +1 212 726 9223 (outside US/Canada). Japan/China/Korea: Nature Publishing Group — Asia-Pacific, Chiyoda Building 5-6th Floor, 2-37 Ichigaya Tamachi, Shinjuku-ku, Tokyo, 162-0843, Japan. Tel: +81 3 3267 8751.

#### CUSTOMER SERVICES

Feedback@nature.com  
Copyright © 2015 Nature Publishing Group

# THE HIJACKED BRAIN

Addiction is a devastating disease that alters the brain’s circuitry, notably in young adults. But the changes need not be permanent: improved understanding of them will help in developing ways to lessen the burden. By Margaret Munro. See a Nature Video at [go.nature.com/e1gqkk](http://go.nature.com/e1gqkk).

## ADDICTION CYCLE

Addiction is characterized by bingeing, withdrawal, cravings, dysfunctional emotions and an inability to abstain from the substance or behaviour. It quickly leads to disruption of interconnected brain circuits that are involved in reward, learning and control.

With time, other brain areas are recruited, including those involved in stress and anxiety — in effect an ‘anti-reward’ system.

**C. Thalamus**  
Acts as a hub to relay sensory information and is also important in regulating arousal.

**B. Nucleus accumbens**  
This region, which receives dopamine from the ventral tegmental area, helps to control desire, satiety and inhibition.

**A. Basal ganglia**  
Interconnected regions that are involved in learning, reward and habit formation.

**G. Hippocampus**  
Important for consolidation of memory.

**F. Frontal cortex**  
Responsible for thoughts and actions. The orbitofrontal cortex is thought to play a part in controlling behaviour.

**E. Amygdala**  
Associated with memory and with emotions, notably anxiety and fear.

### STAGE 1 BINGE/ INTOXICATION

Overstimulation of the reward circuit leads to loss of control and bingeing.

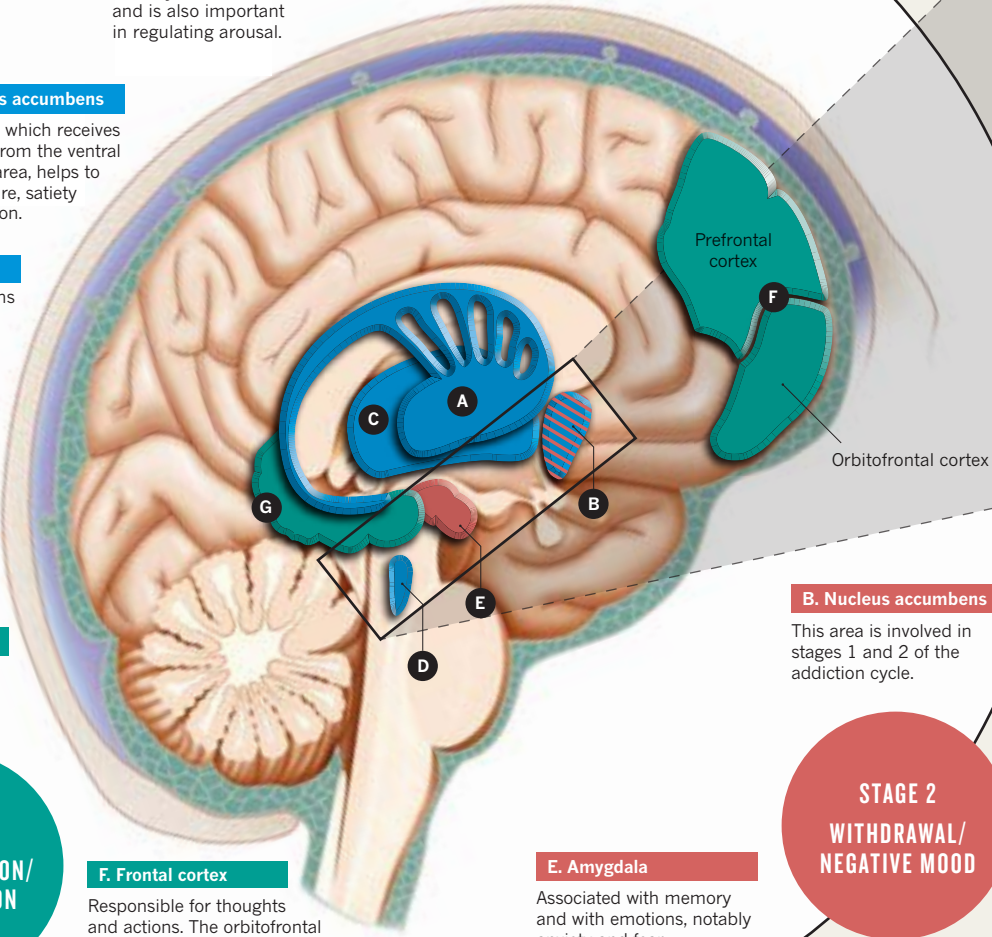
**D. Ventral tegmental area**  
A primitive structure at the top of the brain stem in which dopamine is synthesized (see ‘The dopamine connection’).

### STAGE 2 WITHDRAWAL/ NEGATIVE MOOD

Chronic exposure to addictive substances and behaviours reduces the number of dopamine receptors in the nucleus accumbens (B), so more of the addictive substance or behaviour is needed to feel normal. Changes to the amygdala (E) circuitry have been tied to the irritability, anxiety and stress associated with withdrawal<sup>1</sup>.

People with dependencies have compulsive cravings to repeat the addictive behaviour. Drug misuse is thought<sup>2</sup> to alter the function of the frontal cortex (F) and the hippocampus (G) and helps to embed desires even if they have negative consequences.

### STAGE 3 PREOCCUPATION/ ANTICIPATION



27 MILLION

people had problematic drug use<sup>3</sup> in 2012.

183,000

drug-related deaths were reported<sup>3</sup> in 2012.

1 BILLION

or more people smoke, with the majority living in low- to middle-income countries<sup>4</sup>.

6 MILLION

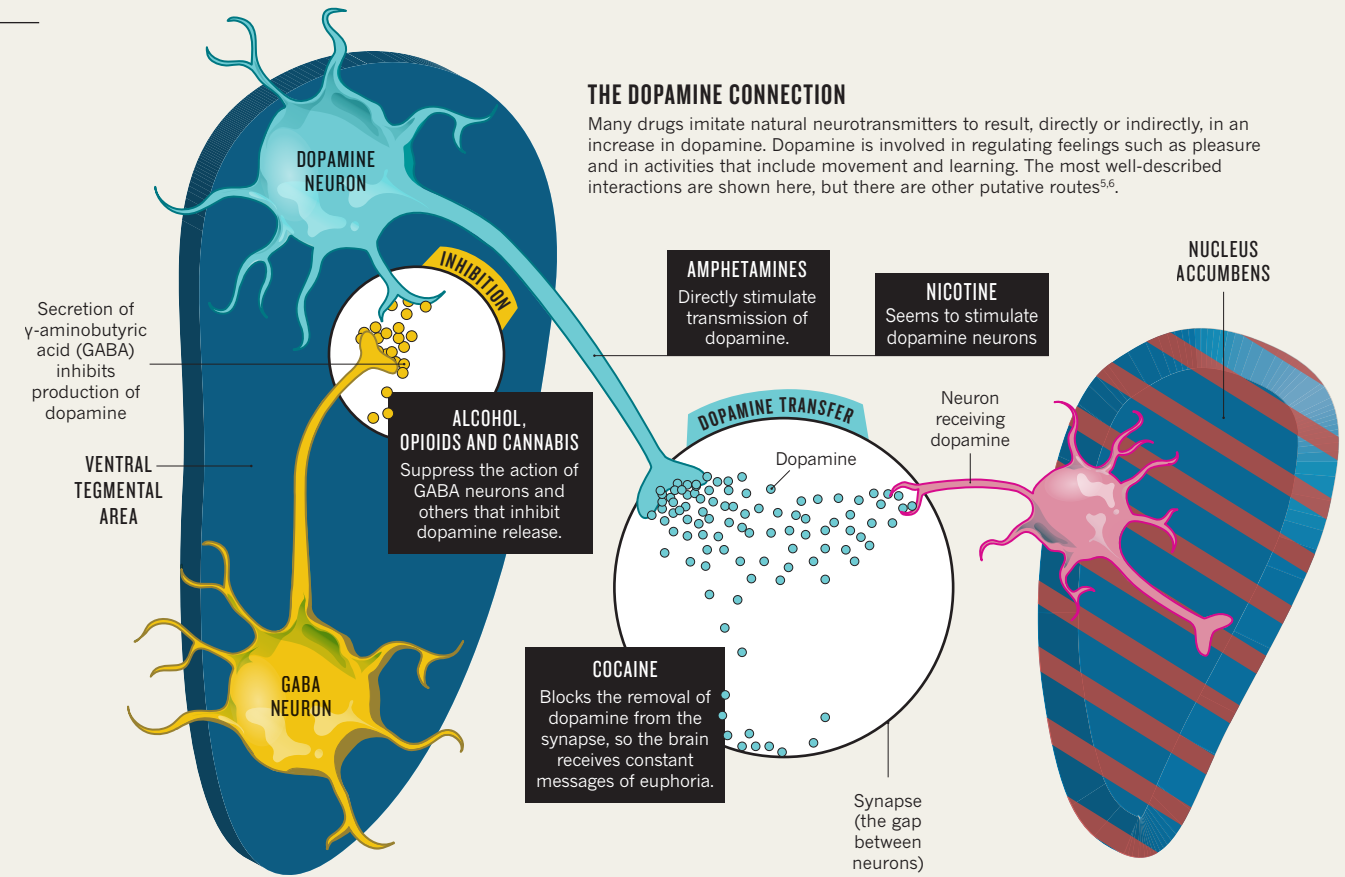
smokers die every year; more than 5 million of the deaths are directly related to tobacco use<sup>4</sup>.

38.3%

of the global population drinks alcohol, with an annual average of 17 litres per person<sup>4</sup>.

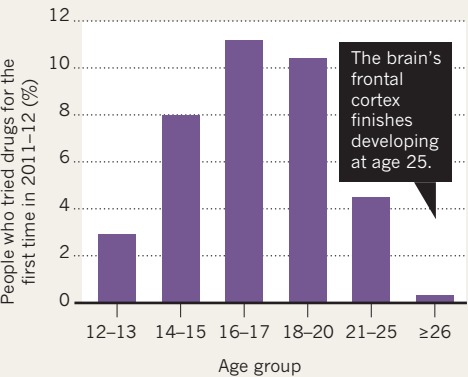
3.3 MILLION

deaths in 2012 were attributed to alcohol consumption<sup>4</sup>.



### DANGEROUS AGE

Many people have their first experience of drugs at a young age, placing them at high risk of addiction. The developing brain may not form properly under the influence of drugs or alcohol<sup>7</sup>.



Sources: 1. Logrip, M. L., Koob, G. F. & Zorrilla, E. P. *CNS Drugs* **25**, 271–287 (2011); 2. Schoenbaum, G. & Shaham, Y. *Biol. Psychiatry* **63**, 256–262 (2008); 3. United Nations Office on Drugs and Crime *World Drug Report 2014* (United Nations, 2014); 4. World Health Organization; 5. Nestler, E. J. *Nature Neurosci.* **8**, 1445–1449 (2005); 6. Fligel, S. B. et al. *Nature* **469**, 53–57 (2011); 7. US National Institute on Drug Abuse.

### HIGH COST OF A HABIT

The estimated annual cost of health care associated with substance misuse in the United States<sup>7</sup>.

ILLICIT DRUGS  
\$11 BILLION

ALCOHOL  
\$25 BILLION

TOBACCO  
US\$130 BILLION





Genetics can influence addictive behaviour later in life, but linking genes to addiction is complicated.

## GENETICS

# No more addictive personality

*The role of temperament, metabolism and development make the inheritance of addiction a complex affair.*

BY MAIA SZALAVITZ

One drunkard begets another, wrote the Greek philosopher Plutarch nearly 2,000 years ago, demonstrating the age-old wisdom of the observation that alcoholism runs in families.

But determining exactly what it is that addicted parents pass down to their children has proved difficult. Scientists have searched for decades for an 'addictive personality' that leaves someone vulnerable to drug problems, but without success. Researchers have tried to identify the genes responsible for addiction, and they have examined the role of early exposure to trauma. Yet they have failed to isolate a single genetic factor that reliably distinguishes between the 10–20% of people who try alcohol or illegal drugs and get hooked and the majority who do not.

Now, however, research into genetics and epigenetics is finally starting to shed some

light on the causes of addiction — and it turns out that the idea of an addictive personality is a myth. Instead, an enormous number of factors, ranging from early life trauma to genes that code for metabolic enzymes, have a role in how the genetics of addiction unfold. By understanding how these factors fit together, researchers hope to develop strategies for the prevention and treatment of addiction.

Plutarch was right to say that addiction is often a familial trait — and it seems that much of this risk is carried genetically. Joni Rutter, director of the Division of Basic Neuroscience and Behavioral Research at the US National Institute on Drug Abuse in Bethesda, Maryland, says that regardless of the drug involved, about 50% of the risk is genetic, within a range of about 40–60%.

Alcoholism is the most widely researched addiction because alcohol

use has such a long history in many cultures. According to George Koob, director of the US National Institute on Alcohol Abuse and Alcoholism in Bethesda, Maryland, the children of people who are dependent on alcohol are 3–5 times more likely to develop the disorder than the rest of the population — and this risk is roughly the same regardless of whether they are raised by their alcohol-dependent parents or adopted by parents who are not dependent on alcohol. The condition is about 60% heritable, he says, adding that this is "reasonably high".

Researchers have managed to demonstrate that genetic predispositions exist, but linking particular genes or traits with addictions has proved much more difficult. Initial genetic findings are often announced with great fanfare, only to fail in replication or be found to have extremely small effects. "Addiction is very heterogeneous," says Rutter, "There are many ways to get there."

## DRUGS AND DISORDERS

Some temperaments and disorders do raise the risk of addiction, however. About half of people with drug-use disorders have an additional psychiatric diagnosis, often a mood, anxiety or personality disorder. "What we're finding is that the addictive personality, if you will, is multifaceted," says Koob. "It doesn't really exist as an entity of its own." Some people with addictions have many personality traits, others have none, but only a few have all of them.

The personality disorder most commonly associated with addiction is antisocial personality disorder (ASPD), which involves dishonest, manipulative, insensitive and criminal behaviour. These characteristics make up the stereotype of someone with an addiction.

"Antisocial behaviour and alcoholism and drug abuse share a bunch of genetic risk factors," says Kenneth Kendler, professor of psychiatry and human genetics at Virginia Commonwealth University in Richmond, who has studied these links in twins. "That's replicated pretty robustly."

A large epidemiological study found that 18% of people with illegal-drug-use disorders have ASPD<sup>1</sup>, as do 9% of people with alcohol-use disorders<sup>2</sup> — much higher than the 4% found in the general population. But although having high levels of antisocial traits is one of the best predictors of substance-use disorders, most people with addictions do not have fully fledged ASPD, and most people with ASPD do not have addictions.

Indeed, many people with substance dependency do not have abnormal levels of antisocial traits at all. However, because breaking the law is itself a diagnostic symptom for antisocial behaviour, this trait will automatically be associated with illegal drug addiction, even if the only laws that are violated are drug laws.

Moreover, being extremely sensitive and overly cautious — essentially the opposite of a callous, impulsive criminal — also raises the risk of addiction, although not by as much.

## ➔ NATURE.COM

A film on a new treatment for addiction is at: [go.nature.com/e1gqkk](http://go.nature.com/e1gqkk)

This suggests both that the stereotype of the addictive personality badly mischaracterizes many people who have a substance-use disorder, and that the genetic risk associated with ASPD does not account for most addictions.

Koob points out that addiction research, like the rest of psychiatry, is increasingly focusing on the genetics that underlie symptoms, such as poor impulse control, rather than on syndromes such as alcoholism or ASPD. “There are specific types of symptoms that have underlying neurobiological bases,” he says.

These temperamental or physiological predispositions can potentially develop into many different disorders. For example, impulsivity could lead to a range of problems: it is a characteristic of addiction, ASPD, bipolar disorder, borderline personality disorder and many more. Impulsive behaviour also increases the risk that teenagers will try drugs — and make it harder for them to resist the urge when they want to stop.

By contrast, anxiety can drive addiction in a different way: people who feel anxious may take drugs to cope with social fears, and their difficulty stopping is not through a lack of control, but because of a lack of alternative ways to manage their emotions. This means that programmes must be tailored to individual needs, not based on the idea that all people with addictions are the same.

## PHYSIOLOGICAL FACTORS

Personality is not the only way in which genes can influence addiction risk. The strongest and most replicated genetic risk factors for alcoholism involve genes linked to metabolism. These genes encode proteins that convert alcohol into acetaldehyde, and acetaldehyde into acetate. Acetaldehyde is particularly toxic, and genes that cause it to build up in the blood, such as a variant of *ALDH2*, make even light drinking unpleasant. “When it is floating in their system, people don’t like it,” says Rutter. “They get really hot or feel nauseous.” Hangovers and the anti-alcohol drug disulfiram produce pretty much the same effect.

Genes that lead to slow alcohol metabolism are common in the Asian population. A 2006 meta-analysis of 15 studies included 4,500 Chinese, Japanese, Korean and Thai participants who were tested for genes related to the metabolism of acetaldehyde and acetate. The largest protective factor was the *ALDH2* variant, which makes people nine times less likely to develop alcoholism than those with other variants of the gene<sup>3</sup>.

But even a gene that provides this much protection can be overridden by environmental pressures. Between 1979 and 1992, for example, the percentage of Japanese people who misused alcohol and who had this variant rose from 2.5% to 13%, as a heavy-drinking culture developed among businessmen that made it much harder to refuse to drink.

One gene that is strongly associated with

cigarette addiction, *CHRNA5*, has essentially the opposite effect on smoking risk as *ALDH2* variant has for alcohol. Having just a single variant can double the risk of nicotine addiction<sup>4</sup>. This link is one of the best supported in any disease, not just in addiction, Rutter says.

Researchers initially thought that the *CHRNA5* variant, which codes for a subunit of the acetylcholine receptor that is affected by nicotine, would make nicotine more pleasurable. This would explain why people who smoke and who have the variant smoke more heavily than those without it. But instead, it softens nicotine’s initial negative effects. Nearly everyone who has ever smoked reports that the first time is nauseating at best. “When I tried cigarettes when I was a kid, I turned green and hated it,” says Rutter.

But people with the *CHRNA5* variant have a less unpleasant experience, says Paul Kenny, a pharmacologist at New York’s Mount Sinai Hospital. “Instead of the drug being more rewarding, what happened was that the aversive effects were diminished,” he says.

Investigation of the *CHRNA5* gene in knockout mice showed that it is active in a brain region called the habenula, which is involved in avoidance and aversion, even though it had not previously been strongly linked with addiction. The evidence also suggests that heavy smoking may damage the habenula by harming the neurons that inhibit it. This would create strong negative feelings and distress in those who smoke, which they may try to fight with even more nicotine.

## ADDICTION AND DEVELOPMENT

Epigenetic mechanisms, which control the activity of genes by switching them on and off, are also being seen as increasingly important in addictions. Kenny’s lab studies these as well and found that one way that addiction can epigenetically ‘rewire’ the brain is by turning on genes that are normally activated only during brain development.

For instance, a mutation in the *MECP2* gene is known to cause Rett syndrome, a developmental disorder found mainly in girls that is associated with intellectual disability and autistic symptoms. During fetal and childhood development, *MECP2* regulates nerve-cell growth, and then it is silenced. However, when rats are allowed to binge on cocaine, *Mecp2* expression “goes through the roof”, says Kenny. Bingeing on cocaine rewires the brain, turning on genes that are usually quiescent in adults.

Other animal experiments have shown that switching off the gene in the reward regions reduces cocaine intake<sup>5</sup>. This suggests that the aberrant learning, which resists the negative consequences of addiction, may be especially deeply engrained. But the actions of *MECP2* have not been studied in normal emotional

learning processes in humans that activate similar circuitry, such as falling in love, so it is not clear whether this is unique to addiction. It is also not known whether these genes are normally reactivated during adolescent brain development, which might help to explain why adolescence and early adulthood is the highest-risk period for addictions. Because Rett syndrome is profoundly disabling, those affected are rarely exposed to drugs, so it is not known how the disorder affects addiction risk.

Kenny thinks that other genes linked to developmental disorders may also be important in addictions — and not just in people who have these conditions, whose brains are wired differently from the start. If addiction does reactivate brain-development genes, more common variants could be involved. “We should be looking for genes that cause developmental disorders,” he says.

Another factor that affects both epigenetics and addiction risk is childhood trauma. Severe stress in early life is known to dramatically increase the risk of addiction, and the risk increases with greater trauma exposure. For example, a recent study of the entire Swedish population showed that people who as children either lost their parents, experienced a parent’s diagnosis of cancer, or witnessed domestic violence had twice the risk of a substance-use disorder later in life compared with those who did not have such stressful experiences<sup>6</sup>.

Indeed, some risk genes, such as those linked to the serotonin transporter, may not cause any problems unless there is a stressful early environment. Both chronic stress and addiction can induce some of the same epigenetic changes in stress systems and in those involved with pleasure, which may partly explain why addiction and trauma are so tightly linked. “Early life experience may dictate whether or not those genes or variations in those genes in those different circumstances tend to come into play,” says Rutter.

Given the increasing evidence of how varied addiction is, treatment and prevention programmes will need to be significantly updated. Some researchers are trying to work out how to target prevention to particular temperaments, rather than attempt to reach both the anxious and the impulsive with the same message.

“That’s the blessing and the curse,” says Rutter. “There are many ways to get there, but that also means many ways to intervene.” ■

**Maia Szalavitz** is a science writer based in New York City.

1. Goldstein, R. B. et al. *Drug Alcohol Depend.* **90**, 145–158 (2007).
2. Goldstein, R. B. et al. *Alcohol. Clin. Exp. Res.* **31**, 814–828 (2007).
3. Luczak, S. E., Glatt, S. J. & Wall, T. J. *Psychol. Bull.* **132**, 607–621 (2006).
4. Fowler, C. D. & Kenny, P. J. *Neuropharmacology* **76**, 533–544 (2014).
5. Im, H.-I. et al. *Nature Neurosci.* **13**, 1120–1127 (2010).
6. Giordano, G. N. et al. *Addiction* **109**, 1119–1127 (2014).

**“Antisocial behaviour and alcoholism and drug abuse share a bunch of genetic risk factors.”**





NIK SPENCER

## NEUROSCIENCE

# Rewiring the brain

*Neuroscientists are learning how to repair neural circuits damaged by addiction.*

BY KATHERINE BOURZAC

Neuroscientist Woody Hopf opens a cabinet in his alcohol research laboratory at the University of San Francisco, California. Inside is a cage containing a rat that is being taught addictive behaviours. The rat has been conditioned to press a lever to release a squirt of alcohol when it hears a beep. Hopf closes the cabinet so that the rat will not be distracted by the sights and sounds of human visitors. Just as it takes time for people to undergo the characteristic brain changes that enforce addiction, he says, it will take time for his rat to become dependent on alcohol.

Researchers such as Hopf view addiction as a disease of the brain circuits responsible for pleasure, stress and decision-making. “Addictive substances come at the brain in different ways,” says George Koob, director of the US National Institute on Alcohol Abuse and Alcoholism (NIAAA) in Bethesda, Maryland. “But in the end, they’re activating some of the same circuitry and patterns of behaviour.”

For decades, researchers have been mapping the electrical and chemical circuits that underlie addiction. Now they are working on strategies for healing these neural pathways. Imaging studies show how the brain rewires during recovery from addiction. When combined with studies of how the brain develops during adolescence, the work could help researchers to understand how the brain changes that are characteristic of addiction occur, as well as who is most

vulnerable and why. This work is rapidly being translated into treatments. By using electrodes and fibre-optic cables, researchers can intervene in neural circuits with great precision, causing animals to lose their taste for alcohol or their interest in cocaine, not just for days but for weeks or months. This work is now being tested in people. Researchers hope that therapies to heal damaged brain circuits will improve the odds of people overcoming addictions.

## CROSSED WIRES

Koob divides addiction into three stages, each with its own brain circuit — groups of neurons or larger structures that interact in a characteristic way (see page S46). Addiction starts with the feel-good binge stage, which is fuelled by the brain’s reward circuit, particularly at the nucleus accumbens. Withdrawal brings stress, centred in the emotional amygdala. Finally, craving and compulsion circuits extending from the prefrontal cortex keep someone using a drug, regardless of negative consequences. Impulsive bingeing leads to habits as the user needs the drug to feel normal.

The changes to the brain’s circuitry are long-lasting, so people trying to give up will often relapse. Even years after recovery, people often start using again when some cue, such as the smell of alcohol or the site of an old hangout, re-triggers old patterns. But the changes are

not permanent. “The brain can enjoy some recovery, probably through remodelling to override the broken parts,” says Edith Sullivan, an experimental psychologist at Stanford University in California.

Some of the physical damage caused by alcohol misuse can be undone. For example, says Sullivan, the brains of people who have misused alcohol for a long period shrink, but some of that brain volume can be regained by sustained sobriety. There is also some functional recovery — even if the pathways are not fully restored, the recovering brain starts to find workarounds.

Sullivan’s group has been using functional magnetic resonance imaging (fMRI) to study cognition in those recovering from alcoholism. A cognitive skill the researchers focused on is spatial working memory — the thinking that helps you to remember where you parked your car, for example. Poor spatial working memory is characteristic of alcohol misuse.

Sullivan’s research suggests that people recovering from alcohol addiction manage to work around brain damage; in other words, their brains find ways of accomplishing tasks by avoiding using damaged areas and they start to regain their working memory<sup>1</sup>. The group found that alcohol-dependent people who had been sober for at least a month performed as well as non-alcohol-dependent controls on spatial working-memory tasks, but used a different part of the brain to do it. Sullivan gave them a more abstract task than looking for a lost object

➔ **NATURE.COM**

A film on a new treatment for addiction is at: [go.nature.com/e1gqkk](http://go.nature.com/e1gqkk)

GARO/PHANIE/REX

or a parked car, but like those tasks it required visual processing, which can take one of two broad neural paths. Patients without brain damage typically rely on a 'where' pathway to do the task, whereas those in recovery from alcohol dependence activate a 'what' pathway, which tends to be used for recognizing and identifying what we see.

"The next step is to find out how to train a person with brain damage to use these new pathways," says Sullivan. Encouraging the natural recovery process could help people who are dependent on alcohol to make faster progress. Sullivan compares the brain damage from alcohol addiction to that caused by stroke. "Recovery won't take three days, it may take three or six months, or a year," she says. It takes time for changes to occur in the brain when someone develops a dependence on alcohol, and it takes time to undo that.

Sullivan is currently investigating whether there is a cost to this rewiring. She suspects that people in recovery are performing the cognitive steps needed for these tasks sequentially, so they take longer than people without addictions who do the steps rapidly in parallel. The damaged brain has fewer circuits to use, so the brain finds it harder to multitask.

## EARLY START

Our understanding of the addicted brain comes from animal studies and from research on people who are already addicted or are in recovery, such as Sullivan's participants. Researchers can only guess at how these changes develop in people. Henning Tiemeier, a psychiatric epidemiologist at Erasmus Medical Center in Rotterdam, the Netherlands, says that the only way to see these changes is to follow people over time. "There is a lot of debate about how harmful substance abuse is for brain development, and you cannot prove it with one brain image," he says.

Two studies, one planned in the United States and one already underway in the Netherlands, could provide some answers. Both will follow adolescents. The adult brain is already formed, although it is still plastic, which is why alcoholism and drug addiction become so engrained, and why the resulting damage cannot be fully repaired. The worry, says Koob, is that the developing brain may not form properly under the influence of drugs and alcohol. Children do not have the cognitive skills to make good choices, making them particularly vulnerable. "Young people have a well-developed reward system but they don't have a good executive control centre," says Koob. The key part of that centre, in the brain's prefrontal cortex, does not finish developing until about the age of 25.

The US National Institutes of Health (NIH), a federal agency that includes the NIAAA and the National Institute on Drug Abuse (NIDA), is currently accepting proposals for the Adolescent Brain Cognitive Development study, which will enrol 10,000 children



A woman receives transcranial magnetic stimulation, a non-invasive therapy that is being used in Italy to treat cocaine addiction.

aged 10 and follow them into adulthood, using neuropsychological tests, brain imaging and surveys, focused specifically on addiction.

Tiemeier is working on the Generation R study in the Netherlands, which has a broader focus on fetal and childhood development and has been following 10,000 children from before birth. The youngest are now aged 9, and the oldest are 12, a stage when some will begin experimenting with cigarettes and alcohol.

Generation R is collecting the first set of brain MRI scans from children in the study, and has about 3,300 so far. By continuing to collect them as the children grow, changes over time will become clear. This is by far the largest brain-imaging study on adolescents in the world, says Tiemeier, so it should provide evidence about how substance use affects the developing organ. He does not expect to see major developmental changes associated with the occasional substance use likely to be found in Generation R because it is a general population study, rather than being focused on people who are addicted to a substance. For this reason, such studies need to be as large as possible if they are to find out what damage drug use does, and how it interacts with puberty, when surges of hormones affect behaviour and brain development.

More information will be available when the Generation R data are combined with results from the NIH study, says Nora Volkow, director of NIDA. These studies will provide a better understanding of the brain changes that reflect what she calls "the skeleton of addictive behaviours". Addiction to cigarettes is different from addiction to heroin, for example, but all addictions have a common neurological framework. These studies will show how it grows. They should also yield insight into who is vulnerable and why, and how they might be helped sooner.

But as further research deepens our understanding of addiction as a disease characterized by changes in the brain, researchers and policymakers need to think about better ways to evaluate medications and therapies, says Volkow. Currently, any pharmaceutical treatment for addiction needs to show that the patient is now completely free of their addiction, which is difficult to prove and takes a long time (see page S53). "Rather than ask for an outcome of complete abstinence, shouldn't we evaluate these treatments on their ability to counteract these brain changes?" she asks.

## PAINFUL REALITIES

This focus on reversing changes to the addicted brain is leading to therapy ideas that are showing promising early results in animals. Hopf's rat studies, for example, have led to a potential therapy for alcoholism that is focused on countering the compulsion to use despite negative consequences such as the loss of relationships with family and friends, employment or health. Because rats do not fear these outcomes, Hopf uses simpler analogues. In some experiments, alcohol-dependent rats are given extremely bitter alcohol instead of the expected normal flavour, or in the lever-pressing test they occasionally receive a painful electric shock to their paw. "The rats want the alcohol but they are not happy about it," Hopf says.

After years of painstaking research and some luck, Hopf found that a particular group of neurons in the reward-centred nucleus accumbens has a key role in promoting compulsive drinking. This year, he found that an approved drug called D-serine binds to receptors on these neurons, causing them to fire less often, leading the

alcohol-dependent rats to drink less<sup>2</sup>. It seems to work by disabling the compulsive behaviour — by turning off the power to deny painful realities. Rats that experience bitter or painful consequences drink less when given the drug. Rats have no such negative consequences to fear and are not affected by the drug and drink as normal.

The nucleus accumbens and a denial of the reality of the situation are involved in multiple stages of addiction, according to Koob, and have a role in both intoxication and the withdrawal process. Hopf is now writing up a plan for a clinical trial of D-serine.

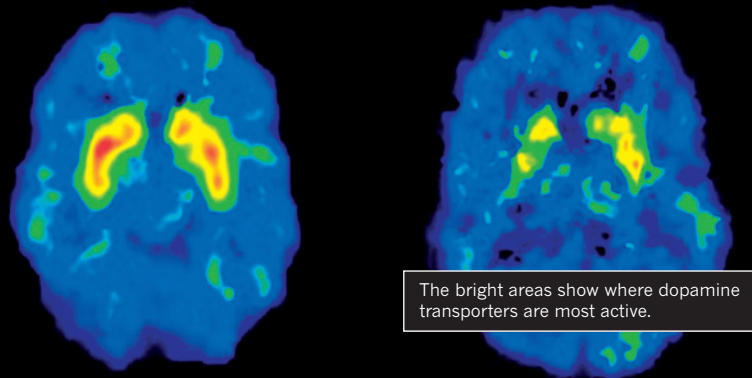
Other techniques target addiction circuits by using physical interventions, rather than drugs. Researchers at the University of Geneva in Switzerland led by neurologist Christian Lüscher have used a method called optogenetics to target a particular group of cells and receptors involved in cocaine addiction in mice. Optogenetics allows researchers to turn off gene expression precisely by shining light into the brain through implanted optical fibres. When Lüscher's group

**"There is a lot of debate about how harmful substance abuse is for brain development."**



## METHAMPHETAMINE MISUSE

Brain scans of someone who has never tried the stimulant methamphetamine (left) and of a user (right). Use of the illegal drug affects dopamine transporters in the brain. Eventually, users need more drug to achieve the desired effect. Dopamine is a neurotransmitter that plays a key part in the brain's pleasure pathways.



used the method to calm a group of overactive dopamine-receptor neurons in the nucleus accumbens, the mice stopped seeking cocaine<sup>3</sup>.

However, optogenetics cannot be used to treat people. The method first requires genetic engineering to render the target cells sensitive to light, and it is not yet possible to safely implant optical fibres in the human brain.

### STIMULATING RECOVERY

Instead, Lüscher's team is attempting to emulate the effects of optogenetics by using methods that translate better to the clinic. They are developing a variation on deep-brain stimulation (DBS), a technique that uses an electric current to silence overactive neurons, which is commonly used to treat movement disorders such as Parkinson's disease. By careful placement of the electrodes, clinicians can target DBS to a particular region in the brain. Researchers have tried using it to treat addiction in people, but results have been mixed.

Lüscher is combining DBS with drugs to block particular receptors in the rat brain, making it possible to silence specific cell types. First they implant an electrode in the nucleus accumbens. Then they use a drug that blocks the neurons' dopamine receptors. Finally, they switch on the electrode for ten minutes. The effects of DBS for treating Parkinson's are transient: when the electric field is turned off, the tremor returns. But Lüscher's combined therapy had a longer-lasting effect<sup>4</sup>. After 10 minutes of stimulation, the rats exhibited normal behaviour for the following 21 days. Lüscher thinks this means that the treatment may be repairing part of the circuit that was damaged by addiction. He says that the group's next step will be to test this approach in primates, or possibly take it to clinical trials.

This demonstration of an apparently long-term reversal of drug-related behaviour is "a miracle", says Jessica Wilden, a neurosurgeon at the Louisiana State University Health Center in Shreveport. Could this lead to a therapy in which you give a patient a pill and a day of brain stimulation and then they are drug free? "In a

small way that's what they're showing," she says. But doing it in people will be harder, she warns.

Wilden is investigating whether DBS can be used to treat methamphetamine (meth) addiction. Meth affects dopamine receptors (see 'Methamphetamine misuse') and is a growing problem, particularly in Iran and in the southern United States, often for military veterans. Unlike other drugs, which tend to be misused mostly by men, meth use is equally common in women, and has a burden on children because women tend to be the primary caretakers, says Wilden.

"I'm trying to set up a stable model of meth abuse, abstinence and relapse in rats, and then try DBS treatment," says Wilden. It is a huge challenge. The drug is a potent stimulant, with effects lasting for 16–20 hours in the rats; the animals become agitated and stressed, and get tangled up in the equipment used to administer the drug and the cables that connect them to the DBS system.

Although DBS is a helpful research tool, Wilden and Lüscher both doubt whether it can be widely used to treat addiction — and Wilden's work with meth illustrates the difficulties. The therapy is expensive, invasive and requires patients to care for the implants and to return to the clinic for regular follow-ups. Those motivated to overcome alcoholism might be able to do it. But people with more destructive addictions, particularly to meth, are less cooperative and have high rates of homelessness, making the treatment even less suitable. "The deep-brain stimulator is a pacemaker, with wires going under the skin into the chest where they connect to a battery," says Wilden. "That's a lot of metal, especially in people who are fragile. There's no way I can implant this in someone living on the streets."

Lüscher and Wilden plan to validate their interventions with optogenetics and DBS in

*"It's up to us to figure out who's getting better and why, and how many sessions it takes."*

animals, and then adapt the results to clinically realistic techniques. The most likely candidate is transcranial magnetic stimulation (TMS), which uses a magnetic field to stimulate electrical activity in neurons deep in the brain. One advantage is that TMS is non-invasive: treatment simply involves wearing a magnetic helmet for a few minutes. It is currently used to treat depression and migraines.

So TMS is more patient friendly, but it is also more mysterious — researchers do not know why it works. Furthermore, it has poor spatial precision, which frustrates neuroscientists who want to target specific brain locations. But this might not matter, says Antonello Bonci, a clinical neurologist and scientific director at NIDA.

In 2013, Bonci published a paper describing how his team had used optogenetics to reactivate an area of the prefrontal cortex that was abnormally quiet in cocaine-addicted rats<sup>5</sup>. The treated rats lost interest in pressing a lever to get cocaine. A few months later, Luigi Gallimberti and Alberto Terraneo at the University of Padova in Italy started using TMS to target the equivalent area in the brains of people addicted to cocaine. They have since been successfully using the technique to treat such people.

Bonci says that the results are anecdotal, but exciting: most people who stuck with the treatment for a few weeks have now been clean for several months, and testify that they do not even think about cocaine any more, he says. With this black cloud lifted, they are able to enjoy food, sex, reading, family time and all the other good things in life. Bonci is now working with the Italian group to design a double-blind clinical trial, and is collaborating with another group to work out how the TMS works. "It's up to us now to figure out who's getting better and why, and how many sessions it takes," he says.

In addition to TMS, the Italian patients also received supportive medical care and psychological therapy. Even with brain stimulation or medication, people still need emotional support, as well as therapy "to identify triggering cues and memories, and practise making new grooves of thought", says Hopf. But with tools such as DBS and TMS, neuroscientists' deepening understanding of the circuitry of addiction is now being translated to the clinic much more rapidly than ever before.

"For the first time in the history of neuroscience, we can think about translating basic science to the clinic in months, as opposed to the 15 years it can take for drug development," says Bonci. Thanks to the new technologies, he says, "we're close to a treatment". ■

*Katherine Bourzac is a science journalist based in San Francisco, California.*

1. Chanraud, S. *et al.* *Cereb. Cortex* **23**, 97–104 (2012).
2. Seif, T. *et al.* *Neuropsychopharmacology* <http://dx.doi.org/10.1038/npp.2015.84> (2015).
3. Pascoli, V. *et al.* *Nature* **509**, 459–464 (2014).
4. Creed, M., Pascoli, V. J. & Lüscher, C. *Science* **347**, 659–664 (2015).
5. Chen, B. T. *et al.* *Nature* **496**, 359–362 (2013).



## PHARMACOTHERAPY

# Quest for the quitting pill

*Addiction researchers are optimistic that they can create effective medication to treat addictions. But the key question is, will pharmaceutical companies bring them to market?*

BY CASSANDRA WILLYARD

In 1948, a group of American businessmen purchased a farmhouse in Center City, Minnesota, and helped to turn the rambling wood structure into a sanatorium for professionals who had become dependent on alcohol. The facility, called Hazelden, spawned one of the largest drug and alcohol addiction treatment networks in the United States, with 16 centres in 9 states.

Since its inception, abstinence has formed the backbone of Hazelden's approach to recovery. But in 2009, Marvin Seppala, the institution's chief medical officer, began pushing for the network to use medication to treat opioid addiction. For the past 20 years the United States has been in the midst of an opioid-addiction epidemic, and as the number of Hazelden residents receiving treatment for opioid dependence grew, Seppala noticed a few disturbing trends. More people seemed to be leaving their programme before completing

their course of treatment, or continuing to use drugs while at Hazelden. Seppala returned to Hazelden in 2009 after two years working in private practice. He had seen the effectiveness of drugs such as Suboxone (buprenorphine and naloxone), an opiate substitute manufactured by Reckitt Benckiser Pharmaceuticals that helps to reduce cravings, and Vivitrol (naltrexone), a long-lasting injectable medication manufactured by Alkermes that blocks the effects of heroin and other opiates. Seppala thought that these medicines might be able to address some of the problems that Hazelden's patients were having adhering to their programmes.

The move was controversial. For decades, Hazelden had helped people with addictions to recover by promoting abstinence and a belief in the power of the 12-step programme, as used by Alcoholics Anonymous. "The use of a maintenance medication like Suboxone wasn't necessarily seen as appropriate," Seppala says. So in 2012, he began holding forums

with Hazelden staff to educate them about his vision. "We thought they were going to throw tomatoes and rotten eggs," he says. But there was surprisingly little resistance. Too many of the clinicians had seen former Hazelden residents relapse and die of a drug overdose.

In 2013, the centre began offering patients Suboxone and Vivitrol as well as group counselling for opioid dependence. Although the number of people involved in the new programme is still small, Seppala has seen some encouraging signs. At Hazelden, the typical dropout rate for people receiving treatment for opioid addiction is 22%, he says. But among those with opiate dependency enrolled in the new programme, the dropout rate was just 5% in 2013 and 2014. Six of Hazelden's patients relapsed and died of opioid overdoses in 2013, but none of them were in the new programme that offers both medication and counselling. "You can't say there's a direct correlation," Seppala says. "However, when it's six to nothing, you've got to say that



there's a dramatic shift, and that we're doing something correct."

Hazleden's adoption of opioid-addiction medication is a sign of a much larger societal shift — a growing recognition that addiction is a complex chronic disease that, like other neurological disorders, often responds to prescription drugs.

But there is nothing on the market for people who are addicted to cocaine or methamphetamines. Only a handful of drugs currently exist to treat nicotine, alcohol and opiate addiction, and those medicines do not always work. "All the addiction medications that are on the market, at best they're successful 30% to 35% of the time," says Stanley Glick, an addiction researcher at Albany Medical College in New York. "We don't only need new medications, we need better medications." Glick and others are exploring a variety of promising targets, and they are optimistic that they can create the next generation of anti-addiction drugs (see 'Drugs against drugs'). But some addiction researchers question whether pharmaceutical companies, which have shied away from addiction therapies in the past, will be willing to bring the advanced therapies to market.

## TARGET PRACTICE

Addictive drugs wreak havoc on the brain's reward circuitry. Some, including heroin, mimic natural neurotransmitters. Others, like cocaine, bind to receptors and prompt the brain to release its own. But the end result is the same: a brain awash in dopamine, the chemical responsible for pleasure. That overlap in the molecular pathways means that it may be possible to develop treatments that target multiple addictions. "What we're interested in is molecular mechanisms that may transcend a particular addictive drug," says Phil Skolnick, director of the division of pharmacotherapies and medical consequences of drug abuse at the US National Institute on Drug Abuse (NIDA) in Bethesda, Maryland. That is important, he says, because "many people who abuse drugs don't abuse just one drug."

Glick thinks that he may have found one such compound. In the late 1980s, Glick received a call from Howard Lotsof, who was formerly addicted to heroin. Lotsof claimed that he had discovered a cure for opiate addiction. He told Glick about a psychoactive compound called ibogaine that occurs in several plant species, including the West African *Tabernanthe iboga* shrub. Lotsof had already approached a number of scientists with his cure. "For better or for worse, I was the first one that was fool enough to become interested in it," Glick says. Glick imagined that he would be able to give the drug to a few morphine-addicted rats and quickly debunk Lotsof's claims. But to Glick's surprise, ibogaine worked. "So we started to get more interested in it," Glick says. Ultimately, it turned out that



*Lobelia inflata* is a source of lobeline, which may help to curb the rush from methamphetamine use.

the drug has some significant drawbacks. It can slow the heart and, at high doses, can damage the nervous system. "There was no way ibogaine was ever going to be an approvable drug in the United States," Glick says.

So Glick partnered with a medicinal chemist and began searching for a new drug, something that would produce the same response as ibogaine, but without all the toxic side effects. The pair landed on a compound<sup>1</sup> called 18-MC. "It doesn't work like any other medication that's ever been proposed to treat addiction," Glick says.

Although some addiction therapies work directly on the circuitry that shuttles dopamine through the brain, the pathway that seems to play a crucial role in most forms of addiction, 18-MC works indirectly. It binds to a nicotinic receptor called  $\alpha$ -3  $\beta$ -4, which is concentrated primarily in the middle of the brain. These receptors are not part of the dopamine pathway, but Glick's research suggests that by blocking the  $\alpha$ -3  $\beta$ -4 receptors, 18-MC dampens the dopamine pathway's euphoric response to drugs<sup>2</sup>. Glick and his colleagues have found that 18-MC works in all kinds of addiction models, curbing animals' use of cocaine, methamphetamines, morphine, alcohol and nicotine. "It opens the door for a whole new approach for affecting the reward system and for reducing addictive behaviour," he says.

Lotsof, who spent much of his life pushing for an anti-addiction therapy, died of cancer in 2010. But Glick kept working to make Lotsof's dream a reality. The same year that Lotsof died, Glick began to work with a biotechnology company called Savant HWP, headquartered in the San Francisco Bay Area, California, to help develop 18-MC further. The first human study began in Brazil in July 2014, led by Savant's South American partner, Brazil-based Hebron Farmaceutica, which is developing

18-MC for a different condition: the parasitic disease leishmaniasis. The collaboration makes good financial sense, Glick says. Both companies need to demonstrate that the compound is safe before they can move forward, and phase I studies, which assess safety in disease-free participants, are similar regardless of the intended use.

The results have yet to be published, but Steven Hurst, Savant's CEO, says that so far, the compound seems to be safe. The next study, slated to begin this year, will start to gather data on whether 18-MC can help people who smoke to break their nicotine habit.

Savant's researchers are not the only ones pursuing the  $\alpha$ -3  $\beta$ -4 receptor as a target for addiction medications. Nurulain Zaveri, a medicinal chemist, was already hunting for medicines to curb nicotine addiction when she learned about Glick's findings in 2003. She was intrigued by the prospect that the largely overlooked receptor could be a good target for nicotine dependence. But she noticed that Glick's compound hit a variety of different targets, not just  $\alpha$ -3  $\beta$ -4. Zaveri wanted something more selective, so she began screening compounds. In 2007, she found one that seemed to be not only selective but also potent — a chemical called AT-1001.

In 2008, Zaveri founded a company called Astraea, headquartered in Mountain View, California, to develop AT-1001 and similar compounds as therapies to help people stop smoking. In 2012, her team showed that AT-1001 can block self-administration of nicotine in rats<sup>3</sup>, and in June 2015, they reported that the compound may also prove valuable for treating alcohol addiction following studies on rats<sup>4</sup>. Zaveri says she also has data to suggest that AT-1001 might help to stop cocaine dependence. Her other leading compounds seem to show similar effects, and Zaveri is currently trying to decide which compound to move into clinical trials.

Linda Dwoskin, an addiction researcher at the University of Kentucky in Lexington,

**"It doesn't work like any other medication that's ever been proposed to treat addiction."**

is working on a different target in the brain's reward pathway. In the 1990s, she began working with lobeline, a compound derived from a group of plants, including *Lobelia inflata*, commonly known as Indian tobacco.

Lobeline binds to nicotinic receptors that are involved in nicotine addiction — others were already investigating it as a potential smoking-cessation tool. But Dwoskin discovered that the compound also binds to a protein in the brain called VMAT2, a transporter that carries neurotransmitters such as dopamine and serotonin. VMAT2 is also the target for methamphetamines, but lobeline did not seem to produce the drug's pleasurable

effects<sup>5</sup>. Dwoskin realized that it might be possible to use lobeline to block VMAT2, thereby preventing the addictive rush associated with methamphetamine use.

Dwoskin launched Yaupon Therapeutics in 2002 and took the compound from the lab to clinical trials. But when people addicted to methamphetamines began taking the lobeline tablets, she immediately realized there was a problem. The drug tasted terrible, and many of the participants developed nausea — not that surprising, because physicians used to prescribe *L. inflata* to induce vomiting, earning it the nickname ‘puke weed’. “It was a minor untoward effect, but enough that compliance to the trial was probably going to be an issue,” Dwoskin says. “We decided we could probably do something better.”

So Dwoskin went back to the drawing board and began working on compounds that would specifically target VMAT2. Over the past decade, she and her colleagues have developed several generations of VMAT2-targeting compounds. “The ones that we’re looking at now are extremely exciting,” she says. They stop animals from self-administering the drug, and even seem to prevent drug-seeking behaviour. “I’ve never seen anything like that before,” she says. Dwoskin will need funding to continue developing the drug in preparation for a human trial. “I feel a need to see this to completion because it looks so promising,” she says.

## A SHOT IN THE DARK

Although many researchers have been focused on developing drug treatments, others have been trying to develop vaccines to curb addiction. The goal is to induce an immune response against addictive substances such as cocaine or nicotine. Then, when the vaccinated individual takes the drug, natural antibodies would prevent the drug’s active ingredient from reaching its target in the brain. Without a pleasurable rush, people might be less prone to relapse. Kim Janda, a chemist at the Scripps Research Institute in La Jolla, California, began working on a vaccine in the 1980s. Over the past three decades, he has worked on vaccines against nearly every type of addictive compound: methamphetamines, cocaine, heroin, nicotine, tetrahydracannabinol (or THC, the active compound in marijuana) and rohypnol. Each one required a different approach.

Of these, Janda thinks that his vaccine against heroin holds the most promise. It combines a heroin-like molecule with a carrier protein designed to elicit an immune response. Heroin breaks down quickly in the body into a compound called 6-acetylmorphine and then into morphine. Janda’s vaccine is designed to mop up all three components, keeping them out of the brain and preventing the rush that heroin typically provides. In 2013, Janda and his colleagues reported<sup>6</sup> that the vaccine seems to prevent both drug-seeking behaviour and relapse in a rat model. In the most challenging

## DRUGS AGAINST DRUGS

A variety of promising pharmaceuticals are currently being developed to treat addiction. But it will be years before any of them join the small number that are already on the market.

| Therapy | Status                  | Developer            | Indication                 |
|---------|-------------------------|----------------------|----------------------------|
| 18-MC   | Phase I clinical trial  | Savant HWP           | Nicotine dependence        |
| AT-1001 | Animal studies          | Astraea Therapeutics | Nicotine dependence        |
| GZ-793A | Animal studies          | Linda Dwoskin        | Methamphetamine dependence |
| HeroVax | Animal studies          | Kim Janda            | Heroin dependence          |
| TV-1380 | Phase II clinical trial | Teva Pharmaceuticals | Cocaine dependence         |

experiment, researchers forced rats that had become addicted to heroin to abstain for 30 days. When they gave the rodents free access again, rats that had received a sham vaccine quickly ramped up their use of the drug, a behaviour that in humans often leads to overdose because the body has lost its tolerance. Vaccinated rats resumed taking the drug, but their consumption did not escalate.

Janda has since tweaked the vaccine and method of injection, and this second vaccine seems to be more effective. But finding someone to help him move to clinical testing might prove difficult. Clinical trials are enormously expensive, and so far Janda has not had much interest from investors or the pharmaceutical industry. He thinks that some companies might also be turned off by previous vaccine failures. A vaccine for nicotine reached a phase III clinical trial in 2009, but ultimately flopped. There is a mentality of “well, you guys had your chance, and it didn’t work”, he says.

## PROFIT MOTIVE

Any compound that makes it into clinical trials risks failure because of unexpected side effects or because it does not work as well as hoped. But some addiction researchers are worried that their experimental therapies will fail for a different reason: lack of interest.

The pharmaceutical industry tends to shun addiction therapies because they are viewed as unprofitable, Janda says. “Pharmaceutical companies don’t view drug addicts as good investments.” But, according to Skolnick, that perception is wrong. In 2012, before Suboxone went off-patent, sales topped US\$1.5 billion. The drug outsold blockbusters like Pfizer’s impotence pill, Viagra (sildenafil). “Those numbers have made it a more interesting game,” Skolnick says. And he thinks that today more companies are willing to take the risk. For example, NIDA recently partnered with Teva Pharmaceuticals, an Israel-based company with the ability to both manufacture and sell medicines, to test the efficacy of a compound called TV-1380 to curb cocaine addiction. Teva “isn’t one of these little biotech companies where once you do a trial, they look for a partner,” he says. “They understood that you can do good and do well at the same time.”

Skolnick thinks there may be an even more serious barrier keeping drug companies at

bay. For addiction therapies, the US Food and Drug Administration views abstinence as the gold standard for approval. That is, the agency wants to see a higher rate of abstinence in the treatment group than the placebo group. So even a medication that helps people to use less of a drug might not gain regulatory approval.

“That seems to be a very, very high bar to jump over,” Skolnick says. “I think that that puts some drug companies off.” Skolnick does not think that such a stringent outcome makes much sense.

“It would seem intuitively obvious, especially for an illegal drug, that if you use it less frequently it would have some benefit.”

To doctors such as Seppala, who witness the aftermath of drug addiction on a daily basis, the need for new medicines seems obvious. “This is a remarkably complex illness,” Seppala says. “Recovery rates are similar to other chronic illnesses, but we don’t feel that’s adequate. Better treatments are necessary.” But Seppala, who has struggled with addiction himself, cautions that even the best medicines will not be a panacea. “People with addiction have often destroyed relationships, done things they don’t even want to admit to anyone,” he says. “If you just give a medication, you’re basically saying it’s only a biological illness and ignoring the rest of this problem.” That is why Hazelden combines a 12-step programme with medication and therapy. To overcome the epidemic of opiate addiction, “we have to use everything at our disposal,” Seppala says. “We can’t rely on a single approach.” ■

**Cassandra Willyard** is a science writer based in Madison, Wisconsin.

- Glick, S. D., Kuehne, M. E., Maisonneuve, I. M., Bandarage, U. K. & Molinari, H. H. *Brain Res.* **719**, 29–35 (1996).
- Taraschenko, O. D., Shulan, J. M., Maisonneuve, I. M. & Glick, S. D. *Synapse* **61**, 547–560 (2007).
- Toll, L. *et al. Neuropsychopharmacology* **37**, 1367–1376 (2012).
- Cipitelli, A., Brunori, G., Gaiolini, K. A., Zaveri, N. T. & Toll, L. *Neuropharmacology* **93**, 229–236 (2015).
- Dwoskin, L. P. & Crooks, P. A. *Biochem. Pharmacol.* **63**, 89–98 (2002).
- Schlosburg, J. E. *et al. Proc. Natl Acad. Sci. USA* **110**, 9036–9041 (2013).



## PERSPECTIVE



# Beyond the neural circuits

To treat addiction, people need help to develop psychosocial skills in addition to taking medication, says **Kenneth E. Leonard**.

We have known many of the fundamental aspects of addiction for some time. Individuals with addictions gradually give up other pleasures in favour of alcohol or other drugs; they experience cognitive difficulties and find it hard to stop thinking about their substance, and they often fail to control its intake or to cope without it. Remarkable progress has been made in identifying neurological factors and changes associated with these behaviours; this in turn has led to pharmaceutical treatments. However, although neurological aspects are crucial, they are not sufficient to explain, prevent or ameliorate addiction.

The development of an addiction is a continuous process, beginning with occasional substance use followed by habitual heavy use and culminating with seemingly incessant or episodically excessive use. Throughout this process, there are neurophysiological changes that occur. After a substantial period of frequent, heavy and patterned use, the term addiction is applied. The concept of disease, representing the underlying neurophysiological changes that occur, is commonly invoked to explain the addictive behaviours. But addressing substance use also requires that we recognize the social and psychological processes that underlie the development of addiction. These processes remain influential even as the circuitry of the brain adapts to the escalating use of a drug. In this sense, addiction is more than a disease and involves more than the brain: it is a systemic behavioural disorder arising from and maintained by psychological, social and biological processes operating both independently and in concert.

In most cases, the development of addiction emerges from an acquired pattern of substance use that serves some motivational function: enhanced positive feelings or states, and reduced unpleasant feelings or states. Individuals who expect that a substance can improve these moods are more inclined to use that substance. Whether someone progresses to addiction depends in part on how well that substance meets those expectations. Over time, sustained excessive substance use can increase negative emotions and reduce the impact of positive ones through the neurophysiological changes to the brain<sup>1</sup>, potentially magnifying the importance of the substance to the individual. But it is important to recognize that increased negative emotions and decreased positive emotions also arise through the impact of use on the individual's cognitive capabilities and the social environment. And that progression into addiction can be impeded by the availability of options other than drugs for enhancing positive states and for coping with negative states.

Furthermore, when it comes to quitting, changes in brain function that have been caused by an addiction do not always lead to cravings, loss of control or relapse. Individuals are able to adjust their intake of a substance or abstain altogether in response to environmental rewards. For example, studies have shown that providing prizes or vouchers for goods or services can increase abstinence from cocaine, tobacco,

alcohol and opiates<sup>2</sup>. Moreover, relapse is influenced by intra-personal processes, including self-efficacy, expectancies about the effects of alcohol, negative emotions and coping abilities. These psychosocial problems remain, even if the underlying neurophysiology of the brain can be rectified.

## A NETWORK OF SUPPORT

An essential element in addressing these issues for each addicted individual is to identify and improve their relationships with friends and family, because these connections directly and indirectly affect both the addictive process and recovery<sup>3</sup>. Addiction makes its mark on a person's social life: he or she may lose friends, alienate family and develop a new social network that supports their habits. The social environment has a significant effect on substance use. Research has demonstrated that marriage and a satisfying marital relationship reduce relapse among men with alcohol dependence and that incorporating relationship counselling into treatment for alcoholism leads to improved outcomes, such as abstinence<sup>4</sup>. Conversely, a critical partner, anger and psychological aggression are associated with poorer outcomes such as relapse<sup>5</sup>. The impact of interpersonal relationships also extends to adult peers and the broader social network, both in terms of general support as well as specific support for abstinence. A network of supportive friends and family are key elements to recovery<sup>6</sup>.

Relieving and perhaps reversing the neurophysiological aspects of addiction with medication is an important element of treatment. However, advances in medication for treatment will not, on their own, cure addiction.

Individuals must also develop the skills and resources to cope with the inevitable negative and stressful experiences of everyday life and to experience meaning and pleasure in social and intimate relationships without the use of the addictive substance. Although these coping processes are represented in the brain, they are most effectively addressed through treatments such as cognitive therapy or training to improve coping skills that focus directly on the psychosocial aspects of addiction. The development and testing of psychosocial strategies to develop these skills and resources must remain an essential aim of research to help prevent and treat addiction. ■

**Kenneth E. Leonard** is the director of the Research Institute on Addictions, University at Buffalo, New York.  
kleonard@ria.buffalo.edu

1. Wise, R. A. & Koob, G. F. *Neuropsychopharmacology* **39**, 254–262 (2014).
2. Benishke, L. A. *et al. Addiction* **109**, 1426–1436 (2014).
3. Hunter-Reel, D., McCrady, B. & Hildebrandt, T. *Addiction* **104**, 1281–1290 (2009).
4. O'Farrell, T. J. & Clements, K. J. *Marital Fam. Ther.* **38**, 122–144 (2012).
5. Mattson, R. E., O'Farrell, T. J., Monson, C. M., Panuzio, J. & Taft, C. T. *J. Fam. Violence* **25**, 33–42 (2010).
6. Stout, R. L., Kelly, J. F., Magill, M. & Pagano, M. E. *J. Stud. Alcohol Drugs* **73**, 489–497 (2012).



CONTINGENCY MANAGEMENT

# Why it pays to quit

*Giving a gift or a cash incentive to someone to give up an addiction sounds like a prize for behaving badly, but the practice works. The real challenge is deciding who should pay for it.*

BY SUJATA GUPTA

Anna smokes almost two packs of cigarettes per day. Her brand of choice is Marlboro, full flavour, no menthol. Anna is 14 weeks pregnant, so that is what her fetus smokes, too.

Aged 33 with short, curly blondish-red hair, and a hot-pink iPhone dangling from the pocket of a hooded sweatshirt, Anna (not her real name) is at the University of Vermont's Center on Behavior and Health in Burlington to learn about a smoking-cessation study for pregnant women. She knows that smoking is bad for her baby. She also knows that she cannot afford the habit, which costs her more than US\$7,000 each year — approximately half of what she earns as an in-home childcare provider.

At the laboratory, Anna learns that she will be randomly assigned to one of two treatment

groups. Treatment A will involve three brief educational sessions and nine phone calls from a trained smoking-cessation coach. At each visit, including the one today, she will earn \$50; completing all 9 calls will earn her an additional \$65. She will also receive money for having ultrasound scans taken of her baby at 30 and 34 weeks and for participating in 8 additional assessments in the year following the birth of her baby. Participants receiving treatment A can be compensated by up to \$655 for their time.

In treatment B, the 'incentives for abstinence' condition, the protocol for participants is the same as for those in treatment A, but they are also screened for drug use through frequent breath and urine tests at the

clinic, and are rewarded when their results are negative. The women earn gift cards that can be used to pay for things such as baby clothes, films, fuel for their car and groceries. In addition to the \$655, participants in treatment B can earn \$1,200–\$2,400 in incentives. "You guys definitely got the right idea, throw money at people," says Anna, who is hoping that she will be assigned to treatment B.

Giving tangible rewards to reinforce positive behaviours, such as abstaining from drugs, underlies a field known as contingency management. Studies<sup>1</sup> have shown that rewards that are contingent on good behaviour help people to refrain from addictive habits, including taking drugs or eating unhealthy food. The rationale is that financial incentives activate the same reward systems in the brain as addictive behaviours.

Moreover, even though quitting an

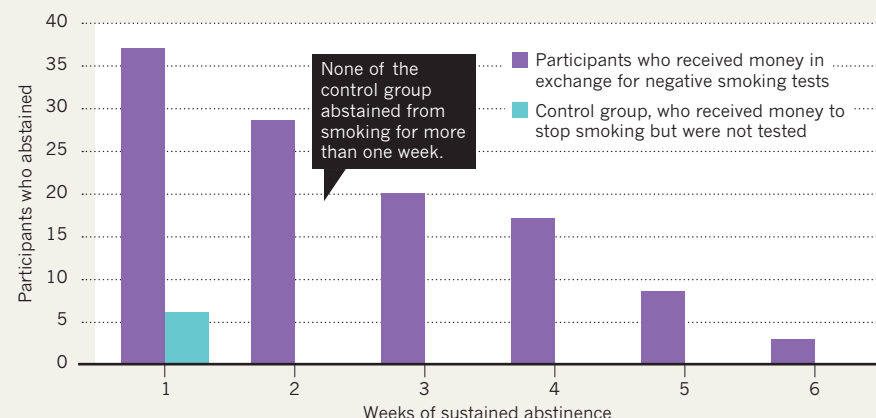
**NATURE.COM**

A film on a new treatment for addiction is at: [go.nature.com/elgqkk](http://go.nature.com/elgqkk)



## SMOKE FREE

In a contingency-management trial, participants who received monetary rewards in exchange for negative smoking tests were more likely to abstain from the habit than those who were paid but were not tested.



addiction improves a person's life in the long run, research has demonstrated that people are biased towards immediate gratification<sup>1</sup>. In other words, receiving an instantaneous cash reward is more desirable than saving the money that would have been spent on drugs and medical care over a long period.

But despite proven results, treatment schemes like this are rarely offered. "I think in some ways that relates to the controversial nature of it," says Nancy Petry, an addiction specialist at the University of Connecticut in Farmington. "You're providing tangible rewards to people who were engaging in illegal or unhealthy behaviours." To help to change that perception, an array of scientists are working on ways to improve contingency management by tweaking how it is offered, in the hope that it will finally cross the boundary to an acceptable and commonplace treatment.

## OFFSHOOT OF THE COCAINE EPIDEMIC

Contingency management arose out of the work of psychologist B. F. Skinner, who challenged the idea that all behaviours stem from free will. Instead, through experiments on rodents that he began in the 1930s, Skinner showed that behaviours can be reinforced by systematic punishments or rewards.

Addiction seemed to be a suitable disease on which to test Skinner's ideas. Addiction is "a concrete behaviour, one that's very difficult to manage, but can be very easily measured", says Maxine Stitzer, an addiction-treatment researcher at Johns Hopkins University in Baltimore, Maryland.

Case studies dating to the 1960s and 1970s provided tantalizing evidence<sup>1</sup> that financial and other incentives might help people to quit drugs or lose weight. But the field really took off in the 1980s with the rise of the cocaine epidemic. Many pharmacological and psychosocial interventions failed, but financial incentives dramatically increased cocaine abstinence.

In the 1970s, Stitzer, one of the first researchers to study contingency management with drug users, and her colleagues developed a simple protocol for people who were addicted to a substance: rewards for negative drug tests and nothing for positive tests. She experimented by providing people who are addicted to opiates with special privileges, such as letting them take methadone (a medication used to control cravings) at home rather than at a clinic. In one study<sup>2</sup>, one group of patients could automatically take methadone home, but another group had to earn the privilege through abstinence from other drugs. Over a four-week period, Stitzer found that people who had to earn their privileges were four times more likely to abstain from drugs (32% abstinence versus 8%) than those who received automatic privileges.

But researchers soon began to favour cash incentives. "Everybody likes money," Stitzer says. "And more importantly, you can vary the amount." Also, unlike Stitzer's methadone approach, which was specific to people who were dependent on opiates, cash incentives can be applied to any addiction. In the 1990s, Stephen Higgins, director of the centre that Anna visited in Vermont and a former postdoc in Stitzer's lab, developed a financial incentive programme for cocaine addiction. Outpatients with negative test results received vouchers that could be redeemed for retail items. By giving vouchers instead of cash, says Higgins, patients can buy something meaningful without using the money for drugs. "What's a reward for you might differ a lot from what's a reward for me," Higgins says. Like cash, "vouchers allow the individual to express what's rewarding for them."

By the 2000s, Stitzer had combined the methadone take-home privilege with

*"Vouchers allow the individual to express what's rewarding for them."*

vouchers. In one experiment<sup>3</sup>, she divided cocaine users on a methadone treatment programme into three groups: one group received take-home methadone privileges for negative test results; a second received the take-home privileges plus vouchers totalling up to \$5,800 over 52 weeks; and a third followed the normal protocol of taking methadone at the clinic. Both sets of people with the take-home privileges submitted three times more drug-negative urine samples than the normal maintenance group over a one-year period. Furthermore, those who received vouchers showed longer periods of abstinence.

A number of meta-analyses<sup>4</sup> have shown that incentives work considerably better than more conventional addiction-management programmes, particularly when vouchers are added to the protocol. What is more, says Higgins, vouchers are the only treatment so far that works for cocaine users.

Contingency management is popular in helping people to stop smoking, an addiction that is seen as more amenable to low-intervention treatment than addictions to other drugs, such as heroin or cocaine. To give an example, in Scotland's deprived Greater Glasgow region, 20% of women report smoking during pregnancy. David Tappin, who specialises in clinical trials for children at the University of Glasgow, recruited more than 600 women, all of whom smoked. At their first prenatal appointment, he randomly assigned half to the routine smoking-cessation programme run by the UK National Health Service (NHS), and the others to the service plus the possibility of receiving up to £400 (US\$610) in shopping vouchers. By the end of their pregnancies, Tappin found that just shy of 9% of women in the control group had stopped smoking, compared to almost 23% in the incentive group<sup>5</sup>.

## REAL-WORLD CHALLENGES

But moving contingency management from the lab to the real world has proved to be challenging. The problems are multifold. For one, asking participants to travel frequently to a drug-testing centre is often untenable. Even though Anna lives only 8 kilometres from the University of Vermont, getting to the lab took her one hour by bus. Then there is the issue of acceptability. To many taxpayers, giving money to people addicted to a drug seems baffling — or morally wrong. But the real hitch is: who should pick up the tab?

Financial incentives make sense in closed health-care systems in which a single entity — such as the NHS — covers all of an individual's health costs, says Petry. Although the initial pricetag may seem high, contingency management is cheaper than treating addiction with more conventional programmes down the line. It can also thwart the sorts of chronic health problems that can arise from prolonged use of a given substance.

Tappin estimates that it would cost the

ALLISON KURTI

NHS £400,000 to roll out a contingency-management programme for all pregnant women who smoke in and around Glasgow. By comparison, he says, the NHS already spends £4.6 million per year on cholesterol-lowering statin drugs in the same region. Compared to other treatments or procedures already in place, “it’s quite a cheap intervention”, he says — and cheaper still if it improves the long-term health of mothers and their babies.

In countries such as the United States, which lack a single-payer health-care system, contingency management can be more difficult to implement. But it can work for specific groups. For instance, the Department of Veterans Affairs, which covers all the health needs of military veterans, adopted a contingency-management programme across its outpatient substance-abuse clinics in 2011 (Petty is an independent consultant for the department). Individuals with negative test results are rewarded with gift cards to the Veteran’s Canteen Service, which sells numerous goods at a discount.

## GETTING CREATIVE

Researchers have also started to experiment with alternative methods to overcome the cost and delivery barriers associated with giving every participant a high-value reward when he or she has a negative drug test.

One way to reduce costs is to encourage people to buy into an incentive programme with their own money, a practice known as a deposit or commitment contract. That money then has to be ‘earned’ back through negative urine tests. In a study published this year<sup>6</sup>, researchers put the concept to the test with more than 2,500 employees of CVS Caremark, a pharmaceutical chain based in Woonsocket, Rhode Island, plus their relatives and friends who wanted to stop smoking.

They divided those individuals into several groups, including one whose members could receive \$800 if they abstained from smoking for six months and another group whose members had to pay \$150 to participate. They could then earn back that \$150 plus a \$650 reward if they did not smoke for 6 months.

The researchers found that both the reward and deposit groups quit at higher rates than those receiving more conventional care, such as nicotine replacement therapies and counselling. Moreover, those who paid the \$150 deposit were twice as likely to quit than those receiving a straightforward reward. But there was a catch. Because participation was optional, 90% of those assigned to the rewards group agreed to give abstinence a try, compared with just 14% in the ‘penalty’ group. “The data are overwhelming in showing the limitation of deposit contracts. The vast majority of folks will not join them,” Higgins says. “The straight incentive programme is the way to go.”

Other solutions have proved to be more successful. For instance, the Department



Vouchers as a reward for a negative drug test.

of Veterans Affairs uses a ‘fishbowl’ model developed by Petty in which participants with negative test results are eligible to draw a slip of paper from a bowl in much the same way as a raffle. Rather than giving everyone the same lump sum, prizes on the slips vary — half of the slips carry no reward, others are for vouchers ranging in value from \$1 to the cost of a new TV. The fishbowl model substantially reduces costs, Petty says, with results rivaling those seen with conventional voucher programmes<sup>7</sup>.

Researchers at Johns Hopkins have developed the ‘therapeutic workplace’, in which jobseekers who abuse substances and present negative drug-test results are rewarded with the opportunity to work and earn wages.

Others are looking at ways to ease the delivery of contingency management. Jesse Dallery, a behaviour analyst at the University of Florida in Gainesville, has developed a model in which people who smoke are sent a carbon monoxide testing kit in the post. Using a web camera or smartphone, participants video themselves blowing into the machine and holding up the value. Negative results are rewarded with vouchers, or by returning a cash deposit made by the participant (see ‘Smoke free’).

Dallery’s former graduate student Bethany Raiff, a behaviour analyst at Rowan University in Glassboro, New Jersey, is now looking to couple that online test with video games — one for smartphones and one for Facebook — in which people who smoke earn rewards in the game instead of money or vouchers. She has partnered with video-game designers who specialize in making health-focused products. “You have to make the games fun and

addictive,” Raiff says — and then corrects her choice of word — “but not in a bad way”.

Each of Raiff’s games lasts about five minutes — the time it takes to smoke a cigarette. To stop people smoking while playing, Raiff and her colleagues are incorporating preventative strategies, such as two-handed playing. In the long-term, Raiff would like to work with insurance companies who have a vested interest in keeping clients healthy. Some, she notes, already cover preventative care, such as gym memberships.

## BACK IN VERMONT

Although contingency management is more effective than conventional treatments, similar issues arise in both approaches — patients struggle to stay abstinent. Bolstered by his short-term success in getting people who have become dependent on cocaine to quit, Higgins began searching for a population in which short-term changes could produce substantial benefits.

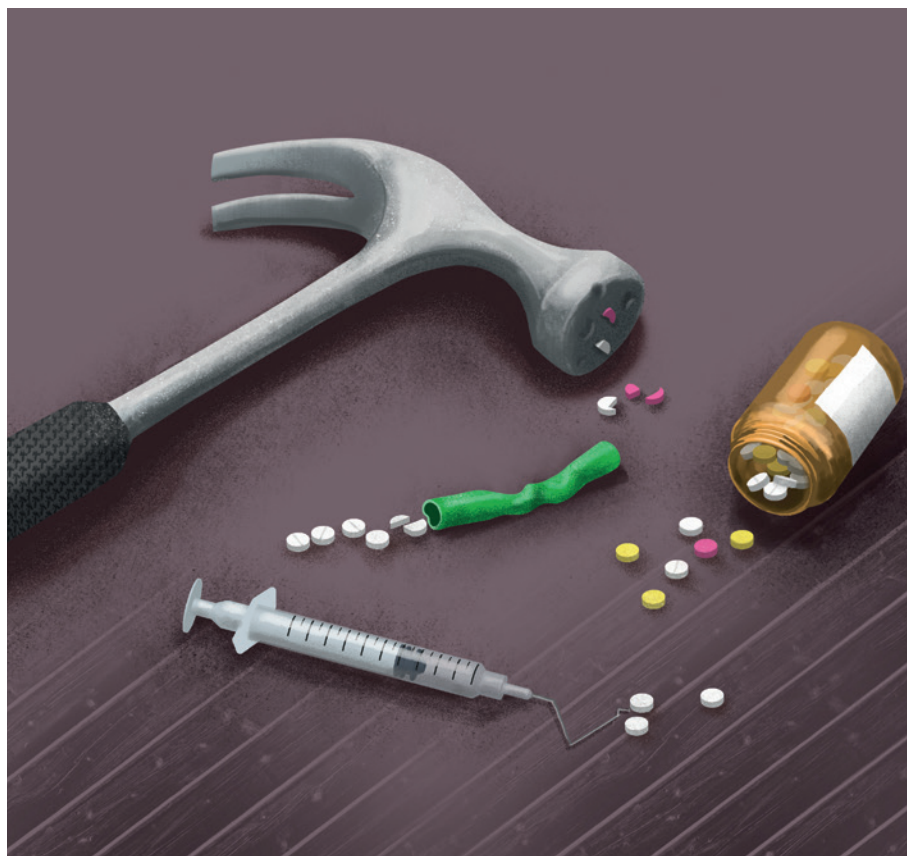
Twelve years ago, Higgins began his research into helping pregnant women who smoke to end their habit. Even short-term abstinence, he reasoned, could dramatically improve their babies’ health. So far, his research has shown that women in the incentive group carry larger fetuses, as measured by ultrasound scans at 30 and 34 weeks, and give birth to larger babies. Now, he is offering incentives to women to continue to abstain from smoking for three months after they give birth, and is considering extending that time frame. His research indicates<sup>8</sup> that providing a reward after childbirth helps women to breastfeed for longer and decreases the risk of postnatal depression.

Back at the centre, Anna finds out that she is in treatment A and will receive compensation for her time but no extra incentives. “There are benefits to both,” Allison Kurti, the post-doc running the study, tells her. “If you’ve got a real busy schedule, it’s kind of nice if we don’t have to see you as often. Some people find that more convenient.” Convenient, perhaps, but as effective? Anna’s not so sure. Just as Skinner observed in rodents so many decades ago, rewards and reinforcement really are key. “I know if someone’s on my ass,” Anna says, “I’m more apt to do it.” ■

**Sujata Gupta** is a freelance science writer in Burlington, Vermont.

- Higgins, S. T., Silverman, K., Sigmon, S. C. & Naito, N. A. *Prev. Med.* **55**, S2–S6 (2012).
- Stitzer, M. L., Iguchi, M. Y. & Felch, L. J. *J. Consult. Clin. Psychol.* **60**, 927–34 (1992).
- Silverman, K., Robles, E., Mudric, T., Bigelow, G. E. & Stitzer, M. L. *J. Consult. Clin. Psychol.* **75**, 839–854 (2004).
- Benishek, L. A. et al. *Addiction* **109**, 1426–1436 (2014).
- Tappin, D. et al. *Br. Med. J.* **350**, h134 (2015).
- Halpern, S. D. et al. *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1414293> (2015).
- Petty, N. M. & Bohn, M. J. *Sci. Pract. Perspect.* **2**, 55–61 (2003).
- Higgins, S. T. et al. *Prev. Med.* **55**, S33–S40 (2012).





in Washington DC. “There are now so many different abuse-deterrent formulations that are either in products or in development that there’s enough variety out there for any product to be able to put abuse-deterrence in it.”

NIK SPENCER

### THE NEW GUARD

Some of the latest tablet formulations are so hard that even a hammer-blow cannot pulverize them. Many pills form a gelatinous goo when dissolved that renders them difficult to inject. Others contain reversal agents that negate the high when the tablets are messed with. The idea is to create pain-relief medicines that are less prone to misuse yet work when taken as directed.

The technologies in place today are not iron-clad, though. A quick perusal of online message boards and videos reveals numerous tips on how to circumvent the defences of even the most reinforced tablets. What is more, not all prescription opioids on the market are misuse-resistant. “We’re still in abuse-deterrent formulations 1.0,” says Richard Dart, director of the Rocky Mountain Poison and Drug Center in Denver, Colorado. But, he adds with a touch of hyperbole, “there are a zillion abuse-deterrent formulations coming”.

Manufacturers have been worried about prescription-drug misuse for decades. When the first controlled-release formulation of the opioid oxycodone hit the US market 20 years ago, the drug’s manufacturer, Purdue Pharma of Stamford, Connecticut, touted the twice-a-day medicine as a less-addictive alternative to the faster-acting painkillers that provide a big opioid hit all at once. In reality, however, Purdue’s longer-lasting pill, sold under the trade name OxyContin, had the opposite effect.

Drug users easily defeated OxyContin’s time-release mechanism by crushing or chewing it. Just one OxyContin could contain more oxycodone than a dozen instant-release pills but no extra ingredients such as paracetamol that make people sick if taken at high doses. OxyContin quickly became the number one addiction problem in many parts of the world, particularly in the United States and Australia. The drug was so popular among the rural poor of Appalachia in West Virginia and Kentucky that it earned the street name ‘hillbilly heroin’.

Purdue set to work to guard against some of the worst forms of misuse. In 2010, the company introduced a misuse-averting version of OxyContin that contains a polymer made of long-chain molecules. This makes the new tablet more difficult to crush — although it is not rock hard. “It behaves more like plastic,” explains Richard Mannion, executive director of pharmaceuticals and analytical development at Purdue. “So, it will deform if subjected to force, but it doesn’t break into a powder easily.” The revised formulation is thus much harder to snort. Plus, Mannion says, when combined with water, the polymer forms a gummy substance that makes it very difficult to draw into a syringe (although misuse is still possible).

### TECHNOLOGY

# Barriers to misuse

*Ingenious pill formulations and the latest manufacturing technologies are helping to stem the tide of painkiller addiction.*

BY ELIE DOLGIN

Mary Marcuccio’s life was turned upside down by drug misuse and addiction. Her son, now 26, started with alcohol and marijuana. Then came cocaine and hallucinogens. By 14, he was stealing prescription painkillers from friends’ medicine cabinets, crushing and snorting the pills to achieve a quick and euphoric high. Within one year, he had graduated to injecting heroin.

This progression is “so stereotypical,” says Marcuccio, founder of My Bottom Line, a Florida-based consulting business for families dealing with substance misuse. According to US survey data, 77% of heroin users say that, like Marcuccio’s son (who remains addicted to heroin), they misused prescription opioids — derivatives of natural or synthetic forms of opium or morphine — before trying heroin.

But substance-misuse specialists think that this chain of addiction might be broken with the aid of the latest manufacturing processes to

make powerful opioid pain medication more resistant to various forms of tampering. Such drug preparations could also save lives. The death toll from misusing prescription opioids has skyrocketed around the world in the past 20 years, with opioid-linked overdoses exceeding fatalities from road accidents or deaths from heroin and cocaine in countries including the United Kingdom, the United States and Australia. “It behooves us to make a greater effort at creating unabusable formularies,” Marcuccio says.

Fortunately, the science and manufacturing of misuse-deterrence are advancing rapidly — and so is the political climate. In the United States — a country that consumes more than 80% of the global opioid supply — politicians are beginning to craft bills to incentivize the development of misuse-resistant formulations. “The idea is to transition the market,” says Dan Cohen, chair of the Abuse Deterrent Coalition, a network of advocacy organizations, technology manufacturers and drug companies based

The new version of OxyContin has proved to reduce the incidence of therapeutic misuse. A study<sup>1</sup> of more than 140,000 people treated at rehabilitation centres across the United States found that misuse by injection, snorting or smoking declined by two-thirds in the two years after the reformulation. In light of these results, in 2013, Purdue won the right from the US Food and Drug Administration (FDA) to describe the misuse-deterrent benefits of OxyContin on the drug's label and to make marketing claims accordingly. The FDA said at the time that any future generic versions of OxyContin would have to incorporate equivalent misuse-deterrent protection. (In April 2015, the FDA released a guidance document outlining the types of study needed to establish misuse-deterrence, but the report stopped short of addressing generic opioid products.)

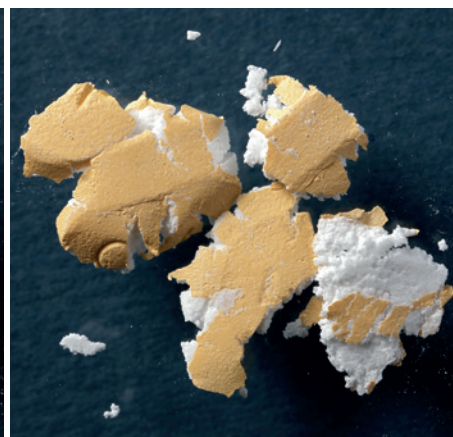
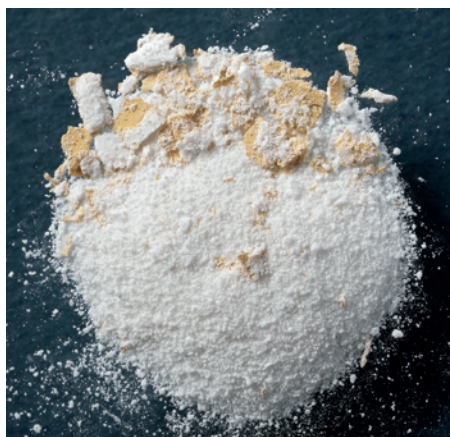
Other painkillers that now have FDA-approved misuse-deterrent labelling include Embeda, an extended-release morphine from New York-based pharmaceutical firm Pfizer, and Targiniq, another long-acting preparation of oxycodone from Purdue. Both contain antagonist agents — offsetting ingredients that remain largely inactive when the drugs are taken as directed, but that will annul the opioid's effects if the drugs are snorted or injected.

"These new technologies are showing some positive results," notes Robert Jamison, a pain psychologist at the Brigham and Women's Hospital Pain Management Center in Chestnut Hill, Massachusetts. In Australia, for example, OxyContin users accounted for more than 60% of the visits to the Medically Supervised Injecting Centre in Sydney. After the tamper-resistant version of OxyContin hit the Australian market in April 2014, a team led by Louisa Degenhardt, a drug-addiction researcher at the University of New South Wales in Sydney, found<sup>2</sup> that the number dropped to 5%. In the United States, levels of opioid misuse have decreased from their peak in 2010, when the new formulation of OxyContin arrived on the market. Rates of opioid dispensing and overdoses have dropped appreciably, too.

These public-health benefits come with an economic bonus. According to calculations from Noam Kirson and his colleagues at Analysis Group, a consulting firm in Boston, Massachusetts, the reformulated OxyContin has reduced misuse-related medical expenses and indirect societal costs by more than US\$1 billion per year in the United States<sup>3</sup>. "These are substantial savings," Kirson says.

## OLD HABITS DIE HARD

Despite the gains, the misuse-deterrence field still has a long way to go. Drug users who have been thwarted by one technology can switch to another prescription medicine that lacks anti-tampering defences. That is what happened in rural Appalachia following the introduction of reformulated OxyContin. Opioid misusers simply started snorting and injecting the less



The original OxyContin pill could be crushed (left) and snorted; and (right) the new tamper-resistant form.

potent immediate-release preparations of oxycodone, most of which lack misuse-deterrence characteristics. "It's kind of a whack-a-mole situation," says Jennifer Havens, an epidemiologist at the University of Kentucky Center for Drug and Alcohol Research in Lexington.

Plus, even with the latest physical defences it is still possible to get high by swallowing lots of OxyContin or Embeda pills at once. Preventing oral misuse requires a different approach — which a company called Signature Therapeutics, based in Palo Alto, California, is pursuing.

Signature Therapeutics' technology uses prodrugs, which are inactive until they undergo the appropriate chemical conversion in the body. When these pills are taken by mouth as directed, a digestive enzyme in the gut called trypsin releases part of the prodrug, initiating the process of opioid drug release. But because trypsin is not found elsewhere in the body, the prodrug remains inert when injected, snorted or smoked. Signature Therapeutics has already tested its painkilling hydromorphone prodrug in a phase I trial of healthy volunteers; the company plans to begin evaluating its oxycodone prodrug in human studies later this year.

Prodrugs alone do not prevent excessive pill-popping, but scientists at Signature Therapeutics have another trick up their sleeves. If the prodrugs look promising in the clinic, the company will add a second compound that blocks trypsin activity. This might seem counterintuitive, but it is all about threshold levels. The amount of trypsin inhibitor found in one or two pills will not interfere with the prodrug modification, but a handful of pills collectively contain enough inhibitor to shut down the conversion process. With this approach, Signature Therapeutics can create either extended-release or immediate-release opioids. Bill Schmidt, chief medical officer at the company, says that the potential of these drugs is "maximum therapeutic benefit with very low abuse liability".

*"It behooves us to make a greater effort at creating unabusable formularies."*

New formulations such as these could ultimately prove to be almost addiction-proof, but they are not cheap. And their benefits might not be fully realized unless authorities require drug companies to include them. "The problem with abuse-deterrence right now is the lack of incentives," Cohen says.

Lawmakers in the US House of Representatives previously proposed legislation that would have barred the approval of any new pharmaceuticals that did not use formulas resistant to tampering. That bill died in committee, but, according to Cohen, revised legislation should be introduced again "soon". Individual US states have also begun to pass laws that compel pharmacists exclusively to dispense, and insurers to cover, misuse-deterrent versions of opioids unless instructed otherwise by a physician.

Ultimately, the success of long-term efforts to rein in opioid addiction could depend on the regulations surrounding generic painkillers. In December 2014, Australia allowed the sale of a generic long-acting oxycodone without misuse-deterrence characteristics. Degenhardt, who is monitoring the drug-misuse data, worries that many of the gains of OxyContin's reformulation will now be lost. By contrast, US authorities have already said that they will not approve such a product.

All of these efforts should help to bring down the number of overdose deaths and also prevent experimentation with prescription pills. In her study population in rural Appalachia, Havens has met so many young people like Marcuccio's son — for whom easily misused opioids were the gateway to addiction — that she has reached a simple, but absolute, conclusion: "The only way that abuse-deterrent formulations are going to work is if they're all abuse-deterrent," she says. "It can't just be piecemeal. It's got to be all or nothing."

**Elie Dolgin** is a science writer in Somerville, Massachusetts.

1. Butler, S. F. *et al.* *J. Pain* **14**, 351–358 (2013).
2. Degenhardt, L. *et al.* *Drug Alcohol Depend.* <http://dx.doi.org/10.1016/j.drugalcdep.2015.02.038> (2015).
3. Kirson, N. Y. *et al.* *Pain Med.* **15**, 1450–1454 (2014).



## PERSPECTIVE



# Behavioural addictions matter

More research, and dedicated funding, is needed to understand and successfully treat compulsive habits, says **Marc Potenza**.

**W**hat behaviours can be considered addictions? Gambling, gaming, Internet use, sex, shopping and eating can become excessive, but whether they should be labelled as addictions is an ongoing debate.

In the most recent, fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)* — a book published by the American Psychiatric Association in 2013 that defines and classifies mental health conditions — gambling disorder was moved from its category of 'Impulse-control disorders not elsewhere classified' to 'Substance-related and addictive disorders'. This represents a significant shift from a view that has prevailed since the 1980s that addictions are disorders involving compulsive drug use, and multiple non-substance-related behaviours may now be considered addictions<sup>1</sup>.

Gambling disorder is currently the only non-substance condition that is listed as an addiction in the *DSM-5*, although a work group proposed that Internet gaming disorder (IGD) warrants additional research. Multiple aspects of IGD remain controversial, including to what extent the Internet may be the vehicle versus the focus of a disorder, and, if a broader 'Internet-use disorder' is to be accepted, the extent to which use represents an addiction. The work group focused on gaming because it was the most well studied and arguably problematic form of Internet use at the time<sup>2</sup>, but behaviours such as social networking and pornography viewing are also under scrutiny. Such uses of the Internet also appear clinically relevant: problematic online social networking, for example, has been linked to poor emotional regulation and problems with alcohol use among university students<sup>3</sup>. Given that more people are growing up with digital technology, considering a broader range of Internet-related activities as potentially addictive seems important for addiction researchers.

## DEFINING BIZARRE BEHAVIOUR

But even if such diagnoses were to be accepted, the question of where to draw the line between abnormal and normal behaviours is still up for debate and has contributed to wide variations in prevalence estimates for problematic Internet use<sup>2</sup>. Currently, the *DSM-5* uses a more stringent threshold for diagnosing gambling disorder (it must meet 4 or more inclusionary criteria out of 9) or Internet gaming disorder (5 or more out of 10) than it does for diagnosing substance-use disorders (2 or more inclusionary criteria out of 11); we must take care not to underestimate how widespread such non-substance behaviours are and the negative impact that they can have on public health.

Another controversial topic is sex addiction. Formal criteria for hypersexual disorder have been proposed and tested<sup>4</sup>, but the condition was not included in the *DSM-5*. As with the other behavioural addictions, debate exists about where to set the threshold between normal and abnormal levels of sexual activity. Nevertheless, similarities in cognitive and biological changes involving craving and reward circuitry

have been noted between compulsive sexual behaviours and substance and gambling addictions, and scales assessing addiction-like features such as craving seem relevant to aspects of sexual behaviours. A better understanding of aetiological and associated factors, such as to what extent psychological and biological determinants linked to gambling and substance addictions also relate to hypersexuality, should help classification efforts and promote the development of targeted treatments.

Other behaviours, including excessive eating and shopping, are also sometimes considered addictions. Of note, patients receiving dopamine-boosting treatment for Parkinson's disease have sometimes developed excessive eating, shopping, sex and gambling habits, suggesting that there may be a biological link that drives all of these behaviours.

But there are nuances: although obesity has been found to share biological features with substance addictions, the variety of ways that the condition manifests itself suggests that only a subset of individuals with obesity may be characterized by food addiction. In particular, individuals with binge-eating disorder are likely to meet food-addiction criteria, suggesting similarities with gambling disorder and substance-use disorders. If foods are demonstrated to have addictive potential, it would be important to identify the specific foods or food components and implement relevant public-health policies and interventions.

While experts debate which non-substance disorders may constitute addictions, people continue to exhibit problematic behaviours. Thus, more research is needed to better understand the epidemiological, clinical, neurobiological,

genetic and cultural features to help prevent and treat behavioural addictions. Research was paramount in compiling the *DSM-5*, and a similar process should be used in writing the 11th edition (due in 2017) of the World Health Organization's *International Classification of Diseases*. But for this to happen, funding agencies must prioritize research into non-substance addictions. In the United States, the National Institutes of Health includes departments focusing on drugs and alcohol, but none that target behavioural addictions. The creation of a national institute on behavioural addictions would help to advance research in this area. In France, the government requires addiction treatment centres to provide care for people with behavioural addictions. Thus, how we classify these behaviours has direct clinical implications, and there is an important need to understand how best to prevent behavioural addictions and help people who experience harm related to them. ■

**Marc Potenza** is director of the Center of Excellence in Gambling Research at Yale University in New Haven, Connecticut.  
e-mail: marc.potenza@yale.edu

1. Potenza, M. N. *Addiction* **101**, 142–151 (2006).
2. Petry, N. M. & O'Brien, C. P. *Addiction* **108**, 1186–1187 (2013).
3. Hornes, J. M., Kearns, B. & Timko, C. A. *Addiction* **109**, 2079–2088 (2014).
4. Reid, R. C. et al. *J. Sex Med.* **9**, 2868–2877 (2012).

*Research into addiction explores many aspects of how and why this disease develops. Here are four of the toughest questions.*

BY DAVID HOLMES

## ADDICTION

# 4 BIG QUESTIONS

### QUESTION

### WHY IT MATTERS

### WHAT WE KNOW

### NEXT STEPS

1

***How broadly should we define addiction?***

We often think about addiction in connection with the misuse of substances such as alcohol, heroin and cocaine. But should compulsive behaviours that involve sex, video games and gambling also be treated as addictions?

Gambling is the only officially recognized behavioural addiction. But there is increasing evidence that other behaviours resemble substance misuse in their underlying neurobiology, and the way in which they respond to treatment (see page S62).

Researchers want to move from a system that classifies addiction according to clinical symptoms to one more rooted in a mechanistic understanding of the disease, with a greater role for genetic imaging and cognitive science.

2

***How is addiction affected by our genes?***

A better understanding of the genetic factors that influence the risk of addiction could revolutionize the way that addiction is diagnosed, treated and prevented.

Numerous adoption and twin studies have shown that the risk of addiction is about 50% heritable. But identifying the genetic factors and the extent to which they interact with each other and the environment has been more tricky (S48).

Researchers are looking in ever-finer detail at potential sources of genetic variations (such as gene copy-number or the existence of rare genes) and epigenetic changes in the way that genes are expressed.

3

***If addiction rewires the brain, can we short the circuit?***

Scientists know that the use of addictive substances causes physical changes in the brain that can lead to addiction. What they are just beginning to understand is how those changes can be reversed to treat the disease.

Using optogenetic-guided brain surgery in mice, researchers have been able to identify a type of dopamine receptor that seems to have a crucial role in addiction. Blocking this receptor has reversed the symptoms of cocaine addiction in mice (S50).

For practical and ethical reasons, optogenetic methods cannot be used in humans, but their increasing use in the lab could speed the discovery of drugs to target and reset the reward circuits that are overloaded in addiction.

4

***Can we get from the 12-step recovery programme to one shot?***

One of the biggest hurdles in treating addiction is preventing relapse: people can fall off the wagon after years of abstinence. A vaccine that neutralizes a substance before it reaches the brain could prevent people from returning to old habits.

Researchers have been developing anti-drug vaccines for more than two decades and have a candidate that tricks rats' immune systems into thinking injected heroin is a pathogen, so the drug is quickly neutralized before it reaches the brain (S53).

Without a trial of an effective vaccine in humans, concerns over whether a vaccine could lead to drug users taking ever-higher doses cannot be addressed, but few, if any, pharmaceutical companies look likely to stump up the cash.

David Holmes is a science journalist based in London, UK.